

ANANTAM IAS • OPTIONAL NOTES

Sociology Optional Research Methods and Analysis

Paper 1 • Unit 3

The most technical unit in Paper 1 and the most reliably scoring one: the qualitative/quantitative divide, every major data collection technique with its real strengths and failures, and the machinery of variables, sampling, hypothesis, reliability and validity that examiners test almost every year.

anantamias.com

(a) Qualitative and quantitative methods. (b) Techniques of data collection. (c) Variables, sampling, hypothesis, reliability and validity.

Qualitative and Quantitative Methods: A Divide of Epistemology, Not Convenience

The commonest mistake is to present this as a choice of tools, numbers versus words. It is not. The divide descends directly from Unit 2: quantitative method is positivism operationalised, qualitative method is interpretivism operationalised. Each carries an implicit theory of what social reality is. Say that in the first line and the rest of the answer writes itself.

- Quantitative method assumes an objective social reality external to the observer, which can be measured, correlated and explained causally. It works deductively, from theory to hypothesis to data. It maximises breadth, large samples, standardised instruments, statistical inference and generalisability. Durkheim's *Suicide* is the archetype.
- Qualitative method assumes social reality is constructed and sustained through meaning, so it must be interpreted rather than measured. It works inductively, generating theory from data. It maximises depth, small numbers, flexible instruments, immersion in natural settings, and validity. Malinowski's *Trobriland ethnography* and Whyte's *Street Corner Society* are the archetypes.
- Clifford Geertz's 'Thick Description: Toward an Interpretive Theory of Culture' (1973), borrowing the phrase from the philosopher Gilbert Ryle, gives the point in one image. A thin description records a contraction of the eyelid. A thick description tells you whether it is a twitch, a wink, a parody of a wink, or a rehearsal of a parody. The physical data are identical; the social facts are entirely different; and only interpretation can separate them. No amount of measurement precision reaches the distinction.
- The emic/etic contrast is a useful compression: qualitative work seeks the emic, the category system of the people studied, in their terms; quantitative work generally works etic, applying the analyst's categories from outside. Blumer's demand to respect the nature of the empirical world is an argument for starting emic.
- Grounded theory, from Barney Glaser and Anselm Strauss' *The Discovery of Grounded Theory* (1967), is qualitative method's most rigorous programme, and it matters because it refutes the charge that qualitative work is unsystematic. Its machinery: theoretical sampling, where the next case is chosen by what the emerging theory needs rather than fixed in advance; constant comparison of incidents against categories; and theoretical saturation as the stopping rule, when new data stop yielding new properties. Note the deliberate inversion of the hypothesis-testing model: theory is the output, not the input.
- Herbert Blumer's 'Sociological Analysis and the Variable' (1956) is the sharpest attack available on the quantitative side, and it belongs in any answer on the limitations of quantitative method. His charges: sociology has no generic variables in the way physics has mass, since 'social class' or 'religiosity' mean different things in different settings; variables are selected arbitrarily; and, decisively, variable analysis leaps from stimulus to response and skips the interpretive process in the middle, which is where the actual social action happens. Correlating religion with suicide tells you nothing about what religion meant to those who died.
- The limitations of qualitative method, which the examiner will want if the question invites balance: small non-representative samples defeat generalisation; the researcher is the instrument, so results are hard to

replicate; analysis is labour-intensive and interpretation is contestable; and there is a real risk of anecdote dressed as evidence.

- The 2022 question on the feminist method points at a distinct qualitative tradition, so name its specifics: rejection of the hierarchical researcher/subject relation in favour of non-exploitative, reciprocal engagement; the use of experience and consciousness-raising as data; Ann Oakley's 'Interviewing Women' (1981), which argued that the standard prescription of a detached, non-committal interviewer is both a masculine ideal and methodologically counterproductive, and that answering respondents' questions and forming a real relationship produced better data; Dorothy Smith's insistence on beginning from the standpoint of women's everyday experience; and Harding's strong objectivity as the epistemological warrant. The distinctiveness is a politics of the research relationship, not a technique.
- The mature position, and the one the 2021 and 2024 questions reward: the divide is real at the level of epistemology but should not be a wall at the level of practice. The methods answer different questions. Quantitative work establishes that a pattern exists and how widespread it is; qualitative work establishes what it means and how it is produced. A study of farmer suicides needs the rate to know there is a problem, and needs the ethnography to know what indebtedness means to a household.
- Mixed method, asked directly in 2024, is the formalisation of that complementarity. Designs: sequential exploratory (qualitative first, to generate categories, then a survey to measure their spread), sequential explanatory (survey first, then interviews to explain anomalies), and concurrent triangulation (both together, converging on the same question). Its strengths: offsets each method's weaknesses, secures breadth and depth, permits triangulation, and produces findings persuasive to more than one audience. Its limitations, which must be stated: it is expensive, slow and demands rival skill sets; results may diverge with no rule for adjudicating between them; in practice the qualitative component is frequently reduced to decorative quotation subordinated to the statistics; and the paradigm incommensurability objection, drawing on Kuhn, asks whether you can coherently combine methods whose underlying assumptions about reality contradict each other. The pragmatist reply is that the research question, not metaphysics, should govern the choice of tools.

Quantitative and qualitative method: the divide runs from epistemology down to technique

Dimension	Quantitative	Qualitative
Epistemological parent	Positivism (Comte, Durkheim)	Interpretivism (Weber, Schutz, Blumer)
Assumed nature of social reality	Objective, external, thing-like, measurable	Constructed and sustained through meaning; must be interpreted
Aim	Explanation of causes; measurement of extent	Understanding of meaning; how conduct is produced
Logic	Deductive: theory to hypothesis to data	Inductive: data to concepts to theory (grounded theory)
Role of theory	Input, to be tested	Output, to be generated
Typical techniques	Survey, structured schedule, experiment, official statistics, content analysis of manifest content	Participant observation, unstructured and focused interview, case study, life history, ethnography
Sample	Large, probability-based, aiming at statistical representativeness	Small, purposive or theoretical, aiming at conceptual richness
Researcher's stance	Detached; instrument standardised to remove the researcher	Immersed; the researcher is the instrument
Standpoint on categories	Etic: the analyst's categories applied from outside	Emic: the members' categories recovered from within
Strong on	Reliability, replication, generalisation, comparison over time	Validity, depth, context, discovery of the unanticipated
Weak on	Validity: imposes categories, misses meaning (Blumer, Cicourel)	Reliability and generalisation: hard to replicate, small N
Archetype	Durkheim, Suicide (1897)	Malinowski, Argonauts (1922); Whyte, Street Corner Society (1943)

KEY TERMS

Thick description — Geertz's term, after Ryle, for an account that renders the layered meaning of an act rather than its physical surface: a wink is not a twitch.

Theoretical saturation — Glaser and Strauss' stopping rule: sampling ends when additional cases yield no new properties of the emerging categories.

Emic / etic — The distinction between an account in the categories of those studied (emic) and one in the analyst's imported categories (etic).

PREVIOUS YEAR QUESTIONS

- What do you understand by 'mixed method'? Discuss its strengths and limitations in social research. [2024]
- How do qualitative and quantitative methods supplement each other in sociological enquiry? [2021]
- What are the different dimensions of qualitative method? Do you think that qualitative methods help to gain a deeper sociological insight? Give reasons for your answer. [2023]
- Analyse the limitations of quantitative methods in social research. [2013]
- Distinguish between quantitative and qualitative techniques of data collection with suitable examples from Indian society. [2018]
- Examine epistemological foundations of qualitative methods of social research. [2017]
- Analyze the importance of qualitative method in social research. [2016]
- What is the distinctiveness of the feminist method of Social research? Comment. [2022]
- How do you view and assess the increasing trend of digital ethnography and use of visual culture in sociological research? [2024]

Observation: Participant and Non-Participant

Observation is the technique where sociology comes closest to anthropology and furthest from the survey. Its promise is unmatched validity: you record what people do, not what they say they do. Its price is everything else.

- Non-participant observation: the researcher watches without joining the group's life. It is less reactive in principle, allows systematic recording against a schedule, and suits public behaviour. Its weakness is that surface conduct is recorded without access to meaning, which is precisely what observation was supposed to deliver. Pure non-participation is also largely a fiction, since a stranger watching a group changes it.
- Participant observation: the researcher takes part in the group's activities over an extended period, sharing its life while recording it. Bronislaw Malinowski established it as a discipline with *Argonauts of the Western Pacific* (1922), studying the kula ring in the Trobriands, and set the standard aim, to grasp the native's point of view, his relation to life, his vision of his world.
- The four classic roles, which give a ready-made structure for any observation answer: complete participant (fully covert, the researcher's identity concealed), participant as observer (a member whose research role is known), observer as participant (mainly observing, with brief involvement, identity known), and complete observer (fully detached and unnoticed).
- Covert versus overt is the axis that decides both the data and the ethics. Covert work avoids the observer effect and gains access to settings that would refuse a researcher, but it means systematic deception, no informed consent, no right of withdrawal, and often the impossibility of taking notes or asking direct questions. Leon Festinger's *When Prophecy Fails* (1956) infiltrated a doomsday cult covertly and produced the theory of cognitive dissonance. Laud Humphreys' *Tearoom Trade* (1970) observed covertly and then traced participants to their homes through their licence plates to interview them under a false pretext; it remains the standard case study in research ethics precisely because the data were valuable and the method indefensible.
- The observer effect, and the Hawthorne effect as its named instance: at the Western Electric Hawthorne Works studies (1924 to 1932), associated with Elton Mayo and reported in Roethlisberger and Dickson's *Management and the Worker* (1939), productivity rose regardless of how illumination and conditions were varied, apparently because workers knew they were being studied. Being observed changes what is observed, which is a problem chemistry does not have.

- Going native, and the linked problem of over-rapport, discussed by S.M. Miller in 'The Participant Observer and Over-Rapport' (1952): the researcher identifies so completely with the group that the analytic distance required to see it collapses, and she becomes an advocate rather than an observer. The opposite failure is remaining so detached that no one confides anything. Competence is the management of that tension, not its resolution.
- Access and gatekeepers: William Foote Whyte's Street Corner Society (1943), a three-year study of an Italian American slum he called Cornerville, entered through 'Doc', and the study is as much a lesson in what a gatekeeper does to a study as in what a street corner gang is. A gatekeeper opens one door and, in doing so, quietly closes others: you inherit his enemies, his blind spots and his position in the group.
- Strengths, stated precisely: high validity and ecological validity, since conduct is observed in its natural setting; access to the unspoken, tacit and routine, which no respondent could report because they have never noticed it; the capacity to discover what the researcher did not think to ask, which is what a questionnaire structurally forbids; and the possibility of studying groups that are hostile, illiterate, or would never complete a form.
- Weaknesses, equally precise: it is time and cost intensive; the sample is one setting, so generalisation is unwarranted; it is not replicable, since the instrument is a person, so reliability is low; recording is selective and memory-dependent; the observer's presence alters the field; and ethical exposure is severe in covert work.
- The 2016 question, which asserts that participant observation is the most effective tool for collecting facts, is inviting disagreement, so disagree with a distinction. It is the most effective tool for collecting meanings, and for facts about conduct in small natural settings. It is a poor tool for facts about distributions, rates and trends in a population, which is what a census or survey exists to supply. Effectiveness is relative to the question, and there is no method that is best in the abstract.
- Digital ethnography, asked in 2024, is participant observation relocated: immersion in online communities, forums, gaming worlds and messaging groups, together with the analysis of visual culture as data. Christine Hine's Virtual Ethnography (2000) is the standard early reference. The genuinely new problems, which are what the question wants: what does participation mean when the researcher is a screen name; what is a field site when the community has no place; consent and the public/private boundary when a forum post is technically public but was written to a small audience; identity play, where the researcher can never be sure who is typing; and the fact that platform architecture, not the researcher, now selects and shapes what is visible.

KEY TERMS

Going native — The failure in which a participant observer's identification with the group destroys the analytic distance needed to study it.

Gatekeeper — The insider who grants a researcher access to a setting, and whose own position within it shapes and limits what the researcher can see.

PREVIOUS YEAR QUESTIONS

- What are the ethical issues that a researcher faces in making use of participant observation as a method of collecting data? Explain. [2023]
- 'Participant observation is the most effective tool for collecting facts.' Comment. [2016]
- Bring out the significance of Ethnography in social research. [2019]

Interview, Questionnaire and Schedule

Three techniques, constantly confused, and the questionnaire/schedule distinction is a favourite because it is a clean discriminator: it separates candidates who have read a methods text from those who are improvising. In India the distinction is not pedantic, it is decisive, because it is the difference between a technique that requires literacy and one that does not.

- Interview types along the structure axis. Structured: fixed questions, fixed order, fixed wording, often fixed response options; high reliability, comparable across respondents, quantifiable; but no probing and no room for the unanticipated. Unstructured: a conversation with a purpose, guided by the respondent's flow; high validity and depth; but low comparability and low reliability. Semi-structured: a topic guide with freedom to probe, which is what most real research uses.
- The focused interview is Robert Merton's, and this is a high-value attribution: Merton, Marjorie Fiske and Patricia Kendall's *The Focused Interview* (1956) developed a technique for interviewing people known to have been exposed to a particular situation, concentrating on their subjective responses to it. The focus group, now ubiquitous in market research and policy work, is the group descendant of Merton's focused interview, a lineage most answers miss.
- The characteristic failures of the interview, and name them as terms: interviewer bias, where the interviewer's tone, wording, expectations or visible identity shapes the answer; social desirability bias, where respondents report the answer that flatters them, which is why self-reported caste prejudice, alcohol consumption and voter turnout are systematically wrong; and the interviewer effect, where the interviewer's own caste, gender, age or class alters what a respondent is willing to say. Ann Oakley's *'Interviewing Women'* (1981) is the standard challenge to the orthodoxy that detachment solves this, arguing detachment produces guarded, impoverished data.
- Focus group discussion, asked in 2022: 6 to 12 participants, a moderator, and a topic. Its distinctive value is that the data are generated by interaction, since participants challenge and refine each other and reveal the socially negotiated character of opinion, which no individual interview shows. Its distinctive dangers are group conformity, dominant speakers monopolising the floor, and moderator influence. The measures the 2022 question asks for: keep the moderator's role facilitative rather than interrogative, use open and non-leading prompts, avoid signalling approval verbally or by expression, actively bring in silent participants and check dominant ones, use homogeneous groups so status differences do not silence people, ensure a neutral venue, record verbatim rather than relying on the moderator's notes, run several groups to expose moderator-specific effects, and have a second analyst code the transcripts.
- The questionnaire: a set of written questions completed by the respondent, mailed, handed over or online. Advantages: cheap; enormous geographical reach; no interviewer bias; anonymity encourages honesty on sensitive matters; the respondent answers at leisure and can consult records; standardised, so reliable and quantifiable. Disadvantages: it requires literacy, which disqualifies it for much of rural India; response rates are low and, worse, non-response is not random, so the returns are biased toward the educated and interested; no probing and no clarification of a misunderstood question; you cannot verify who actually filled it; and unanswered questions cannot be chased.
- The schedule: the same instrument of questions, but filled in by a trained investigator or enumerator in the respondent's presence. Advantages: usable with illiterate respondents, which is why the Census of India and the National Sample Survey use enumerators; response rates are high; the investigator can clarify, probe and observe the setting and the respondent's demeanour, adding observational data the questionnaire can never capture; the identity of the respondent is certain; incomplete answers are chased on the spot. Disadvantages:

expensive, slow, needs trained and supervised field staff; reintroduces interviewer bias and the interviewer effect; and anonymity is lost, so sensitive questions get socially desirable answers.

- The one-line discriminator to memorise: in a questionnaire, the respondent fills the form; in a schedule, the investigator fills it. Everything else follows from that, including the literacy requirement, the cost, the response rate and the bias profile.
- The Indian methodological point worth a sentence in any answer here: given the literacy distribution, the linguistic diversity and the reality that a form arriving by post carries no authority in most households, the schedule administered by an enumerator is not merely one option among several. It is the default for any nationally representative Indian survey, and the mailed questionnaire is largely confined to urban, educated and institutional populations. That is a substantive claim about Indian society, not a technical footnote.
- The 2022 question on information versus data is asking for a real distinction: data are the raw, unprocessed records produced by systematic collection against a purpose, the responses, counts and observations; information is data that has been processed, ordered and given a context so that it answers a question. The subtlety the question flags is that the line is not fixed. What is information at one level becomes data at the next, and no datum is truly raw, since the categories that produced it already embody theory. That is Cicourel's measurement by fiat again: the schedule's response categories decided in advance what could be recorded, so 'raw data' arrives pre-interpreted.

Questionnaire and schedule: the same questions, a different hand holding the pen

Basis	Questionnaire	Schedule
Who fills it	The respondent, personally	The investigator or enumerator, in the respondent's presence
Literacy required	Yes; unusable with non-literate respondents	No; the enumerator reads and records
Cost and speed	Low cost, wide reach, fast	High cost, needs trained field staff, slow
Response rate	Low, and non-response is systematically biased	High; refusals and omissions can be pursued on the spot
Probing and clarification	Impossible; a misread question stays misread	Possible; the investigator clarifies and probes
Interviewer bias	Absent	Present, along with the interviewer effect
Anonymity	High, so better for sensitive topics	Low, so social desirability bias rises
Identity of respondent	Unverifiable	Certain
Supplementary observation	None	The investigator also observes setting and demeanour
Indian usage	Largely urban, educated, institutional populations	The default: Census of India, National Sample Survey

KEY TERMS

Social desirability bias — The tendency of respondents to report the answer that presents them favourably rather than the accurate one, which systematically distorts self-reports on sensitive matters.

Non-response bias — The distortion arising because those who fail to return a questionnaire differ systematically from those who do, so the achieved sample is not the drawn sample.

PREVIOUS YEAR QUESTIONS

- Suggest measures to minimize the influence of the researcher in the process of collecting data through focus group discussion. [2022]
- The difference between information and data in social science is subtle. Comment. [2022]
- Analyse the strengths and weaknesses of social survey method in social research. [2020]
- Discuss the challenges involved in collecting data through census method. [2021]
- Which research technique would be most suitable for the study of consumer behaviour and its social correlates? Explain. [2014]

Case Study, Life History and Content Analysis

Three techniques that share a refusal of the survey's logic: they trade the many for the deep, or they work on traces people left rather than answers they gave.

- The case study is the intensive, holistic examination of a single unit, a person, a family, a village, an institution, an event, in its full context and usually over time. Frederic Le Play's *Les Ouvriers Europeens* (1855), with its intensive family budget studies, is conventionally credited as the method's pioneer in sociology.
- Its logic is not statistical but theoretical. A case study cannot tell you how common something is, and answering that criticism by apologising misunderstands the design. It can tell you that something is possible, which refutes a universal claim with a single case; it can trace a mechanism, showing how a process actually unfolds rather than that two variables correlate; and it can generate the concepts a survey will later measure. Generalisation from a case study is to theory, not to a population.
- Strengths: depth, holism, attention to context and process, high validity, the discovery of the unexpected, and unmatched value in exploratory work and hypothesis generation. Weaknesses: no statistical generalisation, since one village is not India; heavy dependence on the researcher's interpretation; not replicable; and, where a case is chosen because it is interesting, a severe selection problem.
- Life history and biography, asked in 2014: an extended, first-person account of a life, usually gathered over multiple interviews and cross-checked against documents. What it uniquely delivers is exactly Mills' object, the intersection of biography and history, since it shows how large structural changes are actually lived, sequenced and made sense of by a person. It recovers subjective meaning and temporal process, and it gives voice to those written out of official records, which is why it has been central to Dalit, feminist and labour studies. Its dangers are memory's unreliability, retrospective rationalisation, in which people impose a coherent narrative on a life that was not coherent as lived, the self-serving edit, and the impossibility of generalising from one life. The classic exemplar is Thomas and Znaniecki's *The Polish Peasant in Europe and America* (1918 to 1920), built substantially on letters and life records.
- Content analysis works on communication rather than people: newspapers, textbooks, films, advertisements, speeches, social media posts. Bernard Berelson's classic definition, in *Content Analysis in Communication Research* (1952), is the one to quote: a research technique for the objective, systematic and

quantitative description of the manifest content of communication. Klaus Krippendorff's later definition widens it usefully, to a technique for making replicable and valid inferences from texts to the contexts of their use.

- The manifest/latent distinction is the crux. Manifest content is what is literally there and can be counted, the number of times a word appears, the column inches given to a story, the number of female characters. Latent content is the underlying meaning, the ideology, the framing, the silences, and it requires interpretation. Berelson's definition restricts the technique to the manifest, which is what makes it reliable, and which is exactly what critics say makes it thin: counting how often 'Dalit' appears in a newspaper says nothing about how Dalits are framed. Latent coding recovers meaning but reliability falls, so inter-coder reliability must be reported.
- Strengths: it is unobtrusive, since the text cannot react to being studied, which eliminates the observer effect entirely; it permits the study of the past and of the dead; it is cheap, replicable and easily repeated over time, so it is unusually good at trends; and it can handle enormous volumes, which computational text analysis has amplified. Weaknesses: it is confined to what was recorded, and what was recorded is already a selection made by someone with interests; it establishes what was communicated, never how it was received, so inferring audience effects from content is a standard fallacy; and it strips text from the context of its production.
- Unobtrusive measures generally, from Eugene Webb and colleagues' Unobtrusive Measures (1966), are the family this belongs to: physical traces, archives and records, and simple observation, all of which avoid the reactivity that interviews and observation cannot. Their argument is that no method should stand alone, and that multiple operationism, attacking a question with several flawed but differently flawed methods, is the route to confidence.

KEY TERMS

Manifest / latent content — The surface, countable features of a communication as against its underlying meaning; counting the first is reliable but thin, coding the second is rich but less reliable.

Unobtrusive measure — A technique using traces, records or texts that cannot react to being studied, thereby eliminating the observer effect.

PREVIOUS YEAR QUESTIONS

- In what ways biographies could be used to study social life? Explain. [2014]

Variables

A variable is any characteristic that takes more than one value across the units studied. If it does not vary, it is a constant and it can explain nothing. That trivial-sounding sentence is the whole logic of causal analysis: you can only explain a difference by another difference.

- Independent variable: the presumed cause, the one manipulated in an experiment or treated as prior in a survey. Dependent variable: the presumed effect, whose variation is to be explained. The dependent variable depends on the independent one, which is the mnemonic that prevents the commonest slip.
- Intervening variable: the mechanism that stands between cause and effect and through which the effect is transmitted. It is caused by the independent variable and in turn causes the dependent variable. Durkheim's Suicide supplies the cleanest example available and you should use it: religion is the independent variable, suicide rate the dependent, and social integration the intervening variable. Protestantism does not kill

anyone. It produces lower integration, and low integration raises the rate. Naming the intervening variable is what converts a correlation into an explanation.

- Antecedent variable: prior to the independent variable in the causal chain, explaining why the independent variable has the value it does. Extraneous or control variable: not part of the hypothesis but capable of contaminating it, so it must be held constant, matched or statistically controlled. Moderator variable: one that alters the strength or direction of the relationship between independent and dependent, so that the effect exists for one group and not for another.
- The spurious relationship is the concept the whole apparatus exists to defeat: two variables correlate not because either causes the other but because both are caused by a third. The stock illustration is that ice cream sales correlate with drowning deaths, both being caused by summer. In sociology the danger is constant and less obvious, and this is why a bare correlation is never an explanation.
- The elaboration paradigm, developed by Paul Lazarsfeld with Patricia Kendall, is the high-value content here, because it is the systematic procedure for finding out which of these you have. Introduce a test factor into an observed relationship between two variables and see what happens to the original association. It yields four outcomes. Replication: the association persists within every category of the test factor, so the original finding is corroborated. Explanation: the association vanishes and the test factor is antecedent to the independent variable, so the original relationship was spurious. Interpretation: the association vanishes and the test factor is intervening, so the relationship is real and you have just found its mechanism, X to T to Y. Specification: the association holds for one category of the test factor and not another, so you have found the conditions under which the relationship operates.
- Note the elegance and the trap in that scheme: explanation and interpretation are statistically identical, since in both the original association disappears. Only the test factor's temporal position, antecedent or intervening, separates a spurious relationship from a real one with a mechanism, and that is a judgement supplied by the researcher's theory. Statistics cannot decide it for you, which is a compact demonstration that no amount of technique substitutes for theory.
- Operationalisation is where sociological variables are won or lost: converting an abstract concept into an observable, measurable indicator. 'Religiosity' becomes attendance frequency; 'social class' becomes occupation, income and education; 'integration' becomes marital status and church attendance. Every such move is a bet, and it is where validity is decided, because the number you end up with is only as good as the correspondence between concept and indicator.
- The critique that must be present in any 'critically examine' variant, and it is Blumer's, from 1956: sociology has no generic variables, since 'social class' does not mean the same thing in a Bihar village and a Bengaluru office in the way that mass means the same thing everywhere; the selection of variables is arbitrary and driven by what is conveniently measurable; and variable analysis skips the interpretive process by which actors actually construct their conduct, jumping from input to output over the very thing sociology exists to study. Aaron Cicourel's measurement by fiat is the parallel charge.
- The defence, which keeps the answer balanced: without variables there is no comparison, no test, no way of being wrong, and therefore no cumulative knowledge. Variables are how a claim is made accountable to evidence. The correct conclusion is that variable analysis is indispensable for establishing that a pattern exists and how it is conditioned, and insufficient for saying what it means, which is a division of labour, not a defeat.

KEY TERMS

Intervening variable — A variable caused by the independent variable and in turn causing the dependent variable; the mechanism through which an effect is transmitted (Durkheim: integration between religion and suicide rate).

Spurious relationship — A correlation between two variables produced not by any causal link between them but by a third variable causing both.

Elaboration paradigm — Lazarsfeld and Kendall's procedure of introducing a test factor to distinguish replication, explanation (spurious), interpretation (intervening mechanism) and specification (conditional relationship).

PREVIOUS YEAR QUESTIONS

- What is a variable in social research? What are their different types? [2025]
- What are variables? How do they facilitate research? [2023]
- Illustrate with example the significance of variables in sociological research. [2017]
- What are variables? Discuss their role in experimental research. [2015]

Sampling

Sampling exists because populations are large and budgets are not. Its intellectual content is a single claim, which sounds impossible until you see the mechanism: a properly drawn few thousand can speak for a few hundred million. What licenses that claim is randomness, and nothing else.

- Fix the vocabulary first, because these terms are marked. Population or universe: the entire set of units about which you want to conclude. Sampling frame: the actual list from which you draw, which is never identical to the population, and whose defects are inherited silently by every subsequent number. Sampling unit, sample, and sample size. Sampling error: the discrepancy between sample and population arising from chance in the draw, which is calculable only under probability sampling and which shrinks as sample size grows. Non-sampling error: everything else, that is bad frames, non-response, interviewer bias, coding mistakes, and note that it does not shrink with size, so a huge badly-drawn sample is simply a confidently wrong one.
- The dividing line: in probability sampling, every unit has a known, non-zero chance of selection, chance and not the researcher decides, and therefore statistical inference and a calculable margin of error are possible. In non-probability sampling, the chance of selection is unknown, so no inference to the population is warranted and no error margin can be computed.
- Simple random sampling: every unit has an equal chance, drawn by lottery or random number tables. It is the theoretical benchmark and is often impractical, since it requires a complete frame of a possibly enormous population and can scatter the sample across a whole country.
- Systematic sampling: select every k th unit after a random start, where k is the population divided by the sample size. Simple to execute in the field. Its specific danger is periodicity: if the frame has a cyclical order that coincides with k , the sample is systematically skewed, as when every k th house in a colony is a corner plot and corner plots are the wealthy ones.
- Stratified sampling: divide the population into strata that are internally homogeneous and mutually exclusive on a relevant criterion, caste, religion, income, region, then sample randomly within each. Proportionate stratification mirrors each stratum's share of the population; disproportionate oversamples small strata so that they can be analysed at all, then reweights. Its virtue is that it reduces sampling error relative to simple random sampling and guarantees the representation of small groups, which chance might

otherwise miss entirely. In Indian social research this is usually not optional: a simple random sample of a district can easily return too few members of a small community to say anything about them.

- Cluster sampling: sample naturally occurring groups, villages, wards, schools, then study all units within the selected clusters. Its logic is the reverse of stratification: strata should be internally homogeneous, clusters should ideally be internally heterogeneous, each a miniature of the population. It is far cheaper for geographically dispersed populations and does not require a frame of individuals, only of clusters. Its cost is a higher sampling error for a given sample size, because people in one village resemble each other.
- Multi-stage sampling: clusters within clusters, sampled in stages, districts, then villages, then households, then individuals, usually combined with stratification at the higher stages. This is what national surveys actually do: the National Sample Survey uses a stratified multi-stage design, and the point is that the textbook types are building blocks, not rival choices.
- Non-probability types. Convenience or accidental: whoever is available; the mall intercept, the researcher's own students. Cheap and almost worthless for inference. Purposive or judgement: the researcher selects units judged to be informative, which is entirely legitimate when the aim is theoretical rather than statistical. Quota: fill fixed quotas for known population categories, but with the final selection left to the interviewer, which is precisely where it fails, since interviewers approach approachable people. Snowball: existing respondents recruit others, the only realistic route to hidden or stigmatised populations, sex workers, undocumented migrants, drug users, members of a proscribed sect, but the sample follows social networks, so isolates are structurally invisible. Theoretical sampling: Glaser and Strauss' grounded theory procedure, where the next case is dictated by the emerging theory and collection stops at saturation.
- The famous failure to cite for quota sampling: the 1948 US presidential election, where the major polls using quota sampling predicted Dewey would defeat Truman, and Truman won. Quota sampling's flaw is not the quotas; it is that within the quota, selection is left to the interviewer's discretion, so an unmeasurable bias enters at the last step.
- The 2015 question, why random sampling is said to have more reliability and validity, wants the mechanism, not the assertion. Randomness removes the researcher's discretion from selection, which is where systematic bias enters. It ensures that every unit, including those the researcher would never have thought of or would have avoided, has a known chance. It makes the sample's departures from the population random rather than directional, so they cancel rather than accumulate. And, decisively, it permits the calculation of sampling error and confidence intervals, so the study can state how wrong it might be, which no non-probability sample can. But answer with the qualification: randomness guarantees representativeness only in expectation and only in the long run, so any single random sample may still be unrepresentative by chance; it cannot repair a defective sampling frame, and a random draw from a bad list is a bad sample; and it does nothing whatever about non-response, which is where most real surveys actually die. Random sampling secures external validity, which is generalisability, and does not by itself secure measurement validity, since a random sample answering a badly worded question yields precisely representative nonsense.
- When non-probability sampling is the right choice, and say this rather than treating it as second-rate: when no sampling frame exists and cannot be built, which is the normal condition for hidden populations; when the population is rare; when the study is exploratory and the aim is to discover categories rather than measure their spread; and whenever the logic of generalisation is theoretical rather than statistical, as in all qualitative research. A grounded theory study using probability sampling would be committing a category error.

Probability and non-probability sampling: what each buys and what each costs

Type	How units are selected	Chief advantage	Chief limitation	Typical use
Simple random	Equal chance for every unit; lottery or random numbers	The unbiased benchmark; permits full statistical inference	Needs a complete frame; scatters the sample geographically	Small, well-listed populations
Systematic	Every kth unit after a random start	Easy to execute in the field; spreads across the frame	Periodicity in the frame produces systematic skew	Field surveys with an ordered list
Stratified	Random draw within homogeneous strata; proportionate or disproportionate	Lower sampling error; guarantees representation of small groups	Needs prior knowledge of the strata and their sizes	Studies of caste, religion or region in India
Cluster	Sample natural groups, then study units within them	Cheap for dispersed populations; needs no frame of individuals	Higher sampling error, since cluster members resemble each other	Village or school-based studies
Multi-stage	Clusters within clusters, sampled in stages, usually stratified at higher stages	Feasible for national populations with no individual frame	Errors compound at each stage; complex to design and weight	National Sample Survey, Census-linked surveys
Convenience	Whoever is available	Fast and cheap	Unknown and uncorrectable bias; no inference possible	Pilot studies, instrument pre-testing
Purposive	Researcher's judgement of who is informative	Efficient for reaching key informants and rare expertise	Wholly dependent on the researcher's judgement	Qualitative and exploratory studies
Quota	Fixed quotas by category; interviewer picks within the quota	Cheap; matches known population proportions	Selection within quota is at interviewer discretion (the 1948 Dewey/Truman failure)	Opinion and market research
Snowball	Respondents recruit further respondents	The only realistic access to hidden or stigmatised populations	Follows networks, so social isolates are invisible	Sex workers, undocumented migrants, proscribed groups
Theoretical	Next case chosen by what the emerging theory requires; stop at saturation	Builds theory systematically; the stopping rule is principled	No statistical generalisation is claimed or possible	Grounded theory (Glaser and Strauss)

KEY TERMS

Sampling frame — The actual list of units from which a sample is drawn; its defects are inherited by every conclusion, and no randomisation can repair a bad frame.

Sampling error — The chance discrepancy between sample and population, calculable only under probability sampling and reducible by increasing sample size.

Non-sampling error — Error from bad frames, non-response, interviewer bias and coding; it does not shrink as the sample grows, so a large flawed sample is confidently wrong.

PREVIOUS YEAR QUESTIONS

- What is sampling in the context of social research? Discuss different forms of sampling with their relative advantages and disadvantages. [2025]
- Explain the different types of non-probability sampling techniques. Bring out the conditions of their usage with appropriate examples. [2022]
- Explain the probability sampling strategies with examples. [2019]
- Why is random sampling said to have more reliability and validity in research? [2015]

Hypothesis

A hypothesis is a tentative, testable proposition asserting a relationship between two or more variables. Every word in that sentence is doing work: tentative, so it can be abandoned; testable, so it can fail; relationship between variables, so it is not a description.

- William Goode and Paul Hatt, in *Methods in Social Research* (1952), give the definition worth quoting: a proposition which can be put to a test to determine its validity. The emphasis on testability is the whole point, and it connects directly to Popper: a proposition that cannot fail is not a hypothesis, it is a belief.
- Types by function. Null hypothesis (H₀): asserts no relationship or no difference, and it is the one actually subjected to statistical test, because a claim can be disproved decisively but never proved conclusively. Alternative or research hypothesis (H₁): what the researcher actually expects, accepted by implication when the null is rejected. Note the logic and state it, since it is what separates the informed answer: we never prove H₁, we only fail to sustain H₀. This is Popper's asymmetry embedded in statistical practice.
- Types by content: descriptive (asserting the existence, size or distribution of some phenomenon), relational or associational (asserting that two variables co-vary), and causal (asserting that one produces the other). Types by precision: directional, which specifies the direction of the relationship, and non-directional, which asserts only that a relationship exists.
- Type I error is rejecting a null hypothesis that is true, a false positive, claiming a relationship that does not exist. Type II error is failing to reject a null that is false, a false negative, missing a relationship that does exist. The two trade off against each other, and setting a stricter significance level reduces the first while increasing the second.
- Sources of hypotheses, which the 2020 question asks for directly. Theory, which is the most respectable source, since a hypothesis deduced from a theory tests the theory when it is tested. Previous research and its unexplained residues and contradictions. Personal observation and experience. Common sense and folk wisdom, which is a perfectly legitimate source, and this connects Unit 1 to Unit 3: common sense is a good place to get hypotheses and a bad place to get conclusions. Analogy from another discipline, as when Spencer and Durkheim borrowed the organic analogy from biology. Culture and the researcher's social position, which Gouldner would say is always operating whether acknowledged or not.

- Robert Merton's contribution is the one to lead with on sources: hypotheses should be derived from theory, and theories of the middle range, general enough to be theory but specific enough to generate testable propositions, are the bridge between Parsons' grand theory, which generates no testable hypotheses at all, and abstracted empiricism, which tests propositions connected to no theory. Merton's reference group theory and his anomie/strain theory are middle-range theories precisely because they yield hypotheses.
- The qualities of a workable hypothesis: conceptually clear, with terms defined and operationalisable; empirically testable, so it forbids some observable outcome; specific rather than sweeping, since a grand hypothesis cannot be tested; related to a body of theory, so that its confirmation or refutation means something beyond itself; and related to available techniques, since a hypothesis you have no way of investigating is a wish.
- The essential qualification, and the mark of a candidate who has actually read methodology: not all good research begins with a hypothesis, and pretending otherwise contradicts everything in Unit 2. Exploratory studies exist precisely because there is not yet enough known to frame one. Grounded theory explicitly forbids starting with hypotheses, because Glaser and Strauss argued that a prior hypothesis forces the data into categories decided before the researcher understood the setting; there the hypothesis is an output. Ethnography, phenomenology and ethnomethodology do not test hypotheses at all, because their object is meaning, not causal relation. A hypothesis is indispensable to explanatory research, and inappropriate to exploratory and interpretive research. Say which, and why.
- The 2016 question supplies its own example and wants it worked out, so work it: 'illiteracy causes poverty' names two variables and asserts a relation, which makes it a hypothesis. To test it you must operationalise both, illiteracy as inability to read and write with understanding, poverty by an income or consumption threshold, state the direction, check for reverse causation, since poverty also prevents schooling, and control for the obvious third variables, caste, landholding, region, that might be producing both. Notice that the exercise immediately raises spuriousness and reciprocal causation, which is exactly why hypothesis and variable analysis belong in the same answer.

KEY TERMS

Null hypothesis — The proposition of no relationship or no difference; it is what is actually tested, because evidence can decisively refute but never conclusively prove.

Type I / Type II error — Rejecting a true null (false positive) as against failing to reject a false null (false negative); tightening against one loosens against the other.

Middle-range theory — Merton's level of theorising, general enough to be theory yet specific enough to yield testable hypotheses; the bridge between grand theory and abstracted empiricism.

PREVIOUS YEAR QUESTIONS

- What is hypothesis? Critically evaluate the significance of hypothesis in social research. [2025]
- Discuss the importance and source of hypothesis in social research. [2020]
- 'Hypothesis is a statement of the relationship between two or more variables.' Elucidate by giving examples of poverty and illiteracy. [2016]

Reliability, Validity, and Why They Pull Against Each Other

Reliability is consistency: does the instrument give the same result on repetition. Validity is truth: is it measuring what it claims to measure. They are not degrees of the same virtue. They are different virtues, and in sociology they are frequently in direct competition, which is the part the syllabus is really pointing at.

- Reliability is the consistency, stability and repeatability of a measure. If two investigators with the same schedule, or the same investigator on two occasions, get different results, the instrument is unreliable and the data are noise.
- The tests of reliability, which the 2022 question asks for by name. Test-retest: administer the same instrument to the same people after an interval and correlate; the problems are memory of the first administration and real change in the interval, which is indistinguishable from unreliability. Split-half: divide the items into two halves and correlate the scores, which tests internal consistency in a single sitting. Internal consistency measured by Cronbach's alpha, which generalises split-half across all possible splits. Parallel or alternate forms: two equivalent versions of the instrument administered to the same people, which avoids memory but requires that the forms really be equivalent. Inter-observer or inter-coder reliability: independent observers or coders record the same material and their agreement is computed, which is indispensable in observation and in latent content analysis.
- Validity is whether the instrument measures the concept it claims to measure. Its types, in ascending order of rigour. Face validity: it looks right on inspection, which is the weakest form and no evidence at all. Content validity: the items cover the full domain of the concept, so a religiosity scale that asks only about temple attendance has poor content validity, since it omits belief and private practice. Criterion validity: the measure correlates with an external criterion, either concurrently or predictively. Construct validity: the measure behaves as theory says the concept should, correlating with what it ought to correlate with (convergent) and not with what it ought not to (discriminant). Add internal validity, whether the causal inference within the study is sound, and external validity, whether the finding generalises beyond the sample, and ecological validity, whether it holds in natural settings rather than only in the contrived one where it was produced.
- The relationship in one line, and it is asymmetric: reliability is necessary but not sufficient for validity. A valid measure must be reliable, but a reliable measure can be consistently and precisely wrong. A weighing scale set five kilograms high is perfectly reliable, giving the identical answer every time, and entirely invalid. Reliability is about the absence of random error; validity is about the absence of systematic error. That distinction is the answer to most questions in this area.
- Now the trade-off the syllabus is aiming at, which most answers miss entirely. The methods that maximise reliability tend to damage validity, and the reverse. A highly standardised instrument, a fixed-choice schedule, identical wording, fixed order, fixed codes, is reliable precisely because it eliminates discretion: anyone can administer it and anyone can repeat it. But that same standardisation imposes the researcher's categories on respondents whose own categories may be entirely different, forbids probing when an answer is ambiguous, and forces meaningful conduct into pre-set boxes. This is Cicourel's measurement by fiat and Blumer's critique of the variable: precision purchased at the cost of measuring the wrong thing, very consistently.
- The mirror image: immersive qualitative work, long participant observation and unstructured interviewing, achieves high validity because it stays close to actors' own meanings, observes conduct in its natural setting and can pursue whatever emerges. But the instrument is a person, with a particular biography, gender, caste and rapport, and no second researcher could reproduce the study. Reliability, in the conventional sense, is unattainable, and demanding it would destroy the very flexibility that produced the validity.
- State the trade-off as a formula worth remembering: standardisation buys reliability with validity; immersion buys validity with reliability. The instrument that anyone can repeat is the instrument that cannot listen. This is not a defect of sociologists' technique. It follows from the object: meaning is context-dependent, and any instrument standardised enough to be identical across contexts is thereby insensitive to context.

- The 2017 question about resolving reliability and validity in research on inequality is a good test case, so answer it concretely: measure the distribution with reliable standardised instruments, income, landholding, occupation, educational attainment, which give you comparable and repeatable numbers, and then investigate what inequality means and how it is lived and reproduced through ethnography and unstructured interviews, which no schedule can reach. Use the qualitative phase to check whether your indicators are valid, whether caste as coded corresponds to caste as practised, and use the quantitative phase to check whether what the ethnography found holds beyond the village. Neither phase can validate itself.
- Triangulation is the general resolution, and Norman Denzin's *The Research Act* (1970) supplies the four types to name: data triangulation (multiple sources, times, settings), investigator triangulation (multiple researchers, which also supplies inter-observer reliability), theory triangulation (multiple theoretical lenses on the same data), and methodological triangulation (multiple methods on the same question). The underlying logic is Webb's multiple operationism: every method is flawed, but the flaws are different, so convergence across differently flawed methods is genuine evidence, and divergence is itself informative rather than a failure.
- The honest closing qualification: triangulation is not a machine for producing truth. When methods disagree there is no algorithm for deciding which is right, and the assumption that different methods are measuring the same underlying thing is exactly what interpretivists dispute. Treat triangulation as a way of making a finding more defensible and its limits more visible, not as proof.

Reliability against validity: different virtues, in tension

	Reliability	Validity
Question it answers	Does the instrument give the same result on repetition?	Is the instrument measuring what it claims to measure?
Type of error it addresses	Random error	Systematic error
How it is assessed	Test-retest, split-half, Cronbach's alpha, parallel forms, inter-coder agreement	Face, content, criterion (concurrent, predictive), construct (convergent, discriminant), internal, external, ecological
Logical relation	Necessary but not sufficient for validity	Presupposes reliability; a valid measure must be reliable
Favoured by	Standardisation: fixed wording, fixed codes, no discretion	Immersion: flexibility, probing, natural settings, actors' own categories
Method that maximises it	Structured schedule, survey, official statistics	Participant observation, unstructured interview, ethnography
What it costs	Validity: imposes the researcher's categories (measurement by fiat)	Reliability: the researcher is the instrument, so no one can repeat the study
Failure it permits	Being precisely and consistently wrong (the mis-set weighing scale)	Being insightfully right about a case no one else can verify

KEY TERMS

Reliability — The consistency and repeatability of a measure; the absence of random error. Necessary for validity, and by itself no guarantee of it.

Validity — The correspondence between what an instrument measures and the concept it claims to measure; the absence of systematic error.

Measurement by fiat — Cicourel's charge that sociology imposes numerical categories on phenomena whose meaning has not been established, buying precision at the cost of validity.

Triangulation — Denzin's strategy of attacking a question with multiple data sources, investigators, theories or methods, so that differently flawed approaches can converge.

PREVIOUS YEAR QUESTIONS

- What is reliability? Explain the different tests available to social science researchers to establish reliability. [2022]
- How can one resolve the issue of reliability and validity in the context of sociological research on inequality? [2017]
- Why is random sampling said to have more reliability and validity in research? [2015]
- Do you think ethnomethodology helps us in getting reliable and valid data? Justify your answer. [2021]

Last-mile revision

- The qualitative/quantitative divide is epistemological, not a choice of tools: quantitative is positivism operationalised, qualitative is interpretivism operationalised. Say this first.
- Geertz's thick description: the wink and the twitch are physically identical and socially opposite. Measurement cannot reach the difference; interpretation can.
- Grounded theory (Glaser and Strauss, 1967): theoretical sampling, constant comparison, theoretical saturation. Theory is the OUTPUT, not the input.
- Blumer's 'Sociological Analysis and the Variable' (1956): no generic variables in sociology; arbitrary selection; variable analysis skips the interpretive process between stimulus and response.
- Mixed method (2024): sequential exploratory, sequential explanatory, concurrent triangulation. Limitation to name: paradigm incommensurability (via Kuhn), plus cost, skill and the tendency to demote the qualitative part to quotation.
- Participant observation roles: complete participant, participant as observer, observer as participant, complete observer. Malinowski (Argonauts, 1922) founded it; Whyte's Street Corner Society (1943) shows what a gatekeeper does to a study.
- Observer effect, named instance the Hawthorne effect (Western Electric, 1924 to 1932, Mayo; Roethlisberger and Dickson 1939). Opposite failure: going native and over-rapport (S.M. Miller, 1952).
- 'Participant observation is the most effective tool for collecting facts' (2016) invites disagreement: it is the best tool for meanings and for conduct in small natural settings, and a poor tool for rates and distributions. Effectiveness is relative to the question.
- The focused interview is Merton, Fiske and Kendall (1956), and the modern focus group is its group descendant. This attribution is worth marks.
- Questionnaire vs schedule in one line: the respondent fills the questionnaire, the investigator fills the schedule. Everything else follows, including the literacy requirement, cost, response rate and bias profile. Census of India and the NSS use enumerator-administered schedules.
- Content analysis: Berelson (1952), the objective, systematic, quantitative description of the MANIFEST content of communication; Krippendorff widens it to replicable and valid inferences from texts to contexts. Manifest is reliable and thin; latent is rich and needs inter-coder reliability. Content establishes what was communicated, never how it was received.
- Variables: independent (cause), dependent (effect), intervening (mechanism). Durkheim's Suicide is the model: religion (IV), integration (intervening), suicide rate (DV). Naming the intervening variable turns a correlation into an explanation.
- Lazarsfeld and Kendall's elaboration paradigm: introduce a test factor, get four outcomes. Persists throughout = replication. Vanishes, factor antecedent = explanation (spurious). Vanishes, factor intervening = interpretation (mechanism found). Holds in one category only = specification. Explanation and interpretation are statistically identical; only theory about time order decides.
- Probability sampling means a KNOWN non-zero chance of selection, which is what licenses inference and a calculable error margin. Types: simple random, systematic (beware periodicity), stratified (strata internally homogeneous), cluster (clusters ideally internally heterogeneous), multi-stage (what the NSS actually uses).
- Non-probability: convenience, purposive, quota (the 1948 Dewey/Truman failure, because interviewers choose within the quota), snowball (hidden populations; isolates are invisible), theoretical (grounded theory).

- Sampling error shrinks as the sample grows; non-sampling error does not. A huge badly-drawn sample is confidently wrong. No randomisation repairs a defective sampling frame.
- Random sampling secures external validity (generalisability), not measurement validity. A random sample answering a badly worded question yields representative nonsense.
- The null hypothesis is what is actually tested, because claims can be refuted but never conclusively proved. Type I: rejecting a true null (false positive). Type II: failing to reject a false null (false negative).
- Merton: hypotheses should come from theory, and middle-range theory is the bridge between Parsons' grand theory (no testable propositions) and abstracted empiricism (no theory).
- Not all research needs a hypothesis: grounded theory forbids it, and ethnography, phenomenology and ethnomethodology do not test hypotheses at all. Hypotheses are indispensable to explanatory research and inappropriate to exploratory and interpretive research.
- Reliability = consistency = absence of random error. Validity = measuring the right thing = absence of systematic error. Reliability is necessary but not sufficient; the mis-set weighing scale is perfectly reliable and wholly invalid.
- THE trade-off: standardisation buys reliability with validity; immersion buys validity with reliability. The instrument anyone can repeat is the instrument that cannot listen. This follows from the object, meaning is context-dependent, not from sociologists' incompetence.
- Denzin's four triangulations: data, investigator, theory, method. Webb's multiple operationism: every method is flawed, the flaws differ, so convergence is evidence. But triangulation is not a truth machine; when methods disagree there is no rule for adjudicating.

Common errors to avoid

- Presenting qualitative and quantitative as 'words versus numbers' or as a matter of the researcher's taste. They descend from opposed epistemologies (interpretivism and positivism), and an answer that never says so has missed the syllabus point.
- Confusing questionnaire and schedule, or asserting the distinction without drawing the consequences. The examiner wants the implications: literacy requirement, response rate, cost, probing, anonymity, and why India's national surveys must use enumerator-administered schedules.
- Confusing stratified and cluster sampling. Strata are internally HOMOGENEOUS and you sample within every stratum; clusters are ideally internally HETEROGENEOUS miniatures of the population and you sample only some clusters. Reversing this is the single most common sampling error.
- Claiming random sampling guarantees a representative sample. It guarantees representativeness only in expectation and in the long run. Any single random sample can be unrepresentative by chance, and no randomisation repairs a defective sampling frame or fixes non-response.
- Saying a larger sample reduces error, full stop. It reduces sampling error only. Non-sampling error, bad frames, non-response, interviewer bias, coding mistakes, does not shrink with size at all.
- Treating validity as a stronger form of reliability, or listing them as two virtues without noticing they conflict. The syllabus point is the trade-off: standardisation buys one at the cost of the other. An answer that misses the tension misses the question.
- Writing that the alternative hypothesis is 'proved' when results are significant. The null is rejected; H₁ is never proved. State the asymmetry, since it is Popper's logic embedded in statistical practice.
- Confusing Type I and Type II errors. Type I rejects a TRUE null (false positive, claiming a relationship that does not exist); Type II fails to reject a FALSE null (false negative).
- Confusing an intervening variable with an extraneous or antecedent one. An intervening variable is caused BY the independent variable and in turn causes the dependent one; it transmits the effect rather than confounding it. An antecedent variable sits before the independent variable, and a spurious relationship arises when a third variable causes both.
- Asserting that a case study's inability to generalise is a fatal flaw. Its generalisation is to theory, not to a population: it establishes possibility, traces mechanisms and generates concepts. Judging it by survey criteria is a category error, and so is judging ethnomethodology by conventional reliability.
- Defining content analysis without Berelson's restriction to MANIFEST content, and then claiming it reveals ideology. Latent coding is possible but reliability drops and inter-coder agreement must be reported. Also, content analysis never establishes audience reception; inferring effects from content is a standard fallacy.
- Answering the mixed method question with advantages only. The 2024 question says strengths AND limitations, and the paradigm incommensurability objection is the one that separates a good answer from a listing.