ACROSS MARKETS AND BORDERS

Enterprise, Trade and Leadership with Purpose and Integrity

PART I | FOUNDATIONS OF TRUTHFUL BUSINESS INQUIRY

CHAPTER 3

RESEARCH METHODS FOR BUSINESS AND TRADE

Questions, Evidence, Design, Ethics and Scholarly Inquiry

Sixbert SANGWA
Department of International Business and Trade
African Leadership University, Kigali, Rwanda

Open Christian Press
2026

CHAPTER 3 | RESEARCH METHODS FOR BUSINESS AND TRADE

Questions, Evidence, Design, Ethics and Scholarly Inquiry

First edition, 2026

Chapter DOI: https://doi.org/10.65655/ocp.m23.c62

Book DOI: https://doi.org/10.65655/ocp.m23

Author: Sixbert SANGWA

Affiliation: Department of International Business and Trade, African Leadership University, Kigali, Rwanda

Corresponding author: ssangwa@alueducation.com

Publisher: Open Christian Press, a Publishing Imprint of Open Christian University and Center for Faith and Work.

Publisher website: press.openchristian.education

Publisher correspondence: editor@openchristian.education

Suggested citation. Sangwa, S. (2026). Research methods for business and trade: Questions, evidence, design, ethics and scholarly inquiry. In *Across markets and borders: Enterprise, trade and leadership with purpose and integrity*. Open Christian Press. https://doi.org/10.65655/ocp.m23.c62

Abstract

Research methods determine whether business and trade claims deserve confidence. This chapter develops a comprehensive undergraduate foundation for designing rigorous applied research, from the first practical problem to a defensible proposal and dissemination plan. It explains ontology, epistemology, paradigms, theory-evidence relationships, research questions and hypotheses; systematic literature searching and synthesis; qualitative, quantitative, mixed-methods, case-study, survey, experimental, quasi-experimental, documentary and secondary-data designs; sampling and sample-size reasoning; operationalization, reliability, validity, trustworthiness and bias; data collection and analysis planning; research ethics, data protection and organizational permissions; transparency, reproducibility and responsible use of generative AI. African business research is central, including the Nigeria YouWiN! randomized experiment and World Bank Enterprise Surveys methodology, while an Indian management field experiment provides comparative perspective. Christian moral reasoning treats inquiry as accountable truth-seeking that must respect persons, refuse fabrication, represent uncertainty honestly and steward evidence for the good of organizations and communities.

Keywords: business research; trade research; research design; research paradigms; sampling; measurement; qualitative research; mixed methods; secondary data; research ethics

Learning Outcomes

By the end of this chapter, the reader should be able to:

1. Distinguish practical business problems, research problems, research purposes and decision questions, and convert a broad concern into a feasible, researchable problem.
2. Explain ontology, epistemology, axiology and methodology, and compare postpositivist, interpretive-constructivist, critical realist, pragmatic and critical-emancipatory research orientations without reducing them to slogans.
3. Develop aligned research objectives, questions and, where appropriate, hypotheses that are precise, answerable and proportionate to the proposed evidence.
4. Design a transparent literature search, synthesize rather than merely list prior studies, distinguish theoretical and conceptual frameworks, and justify a genuine research gap.
5. Compare quantitative, qualitative, mixed-methods, case-study, survey, experimental, quasi-experimental, longitudinal, cross-sectional, documentary and secondary-data designs, selecting a design that fits the question rather than personal preference.
6. Define target populations, sampling frames and units of analysis; compare probability and non-probability sampling; and justify sample size through precision, statistical power, information power, saturation or case logic as appropriate.
7. Operationalize constructs and assess measurement quality using reliability, validity, trustworthiness, reflexivity and bias controls appropriate to the design.
8. Design defensible surveys, interviews, focus groups, observations, document studies and secondary-data protocols, including piloting and field-quality procedures.
9. Produce an analysis plan that specifies how each question will be answered without prematurely duplicating the statistical techniques taught in Chapter 4 or advanced analytics in Chapter 5.
10. Apply research ethics, informed consent, confidentiality, data protection, organizational permissions and research-integrity principles to business and trade inquiry, including Rwanda's data-protection context and cross-border data responsibilities.
11. Use generative AI in research only under explicit human accountability, verification, privacy and disclosure controls, and identify uses that would fabricate, distort or improperly substitute for scholarship.
12. Build an aligned research proposal, project plan and dissemination strategy, and audit the complete design for methodological, ethical and moral integrity before data collection begins.

Concept Map and Chapter Roadmap

PROBLEM & PURPOSE	PHILOSOPHY & THEORY	QUESTIONS / HYPOTHESES	RESEARCH DESIGN
SAMPLING & MEASUREMENT	EVIDENCE COLLECTION	ANALYSIS PLAN	CLAIMS & DISSEMINATION
ETHICS • INTEGRITY • DATA PROTECTION • ACCOUNTABILITY			

Figure 3.1. Research inquiry as an aligned and ethically bounded system.
Source: Author's synthesis.

Research is not a sequence of isolated forms to complete. It is an alignment problem. A defensible study begins with a question worth answering, adopts assumptions that are coherent with the kind of knowledge sought, selects a design

capable of generating relevant evidence, obtains that evidence through transparent procedures, analyzes it with methods suited to the data and claim, and communicates conclusions no more strongly than the evidence permits. Ethics and integrity are not a final approval box. They surround the entire architecture because the choice of problem, sample, measures, field procedures, analysis and reporting can benefit or harm people and can clarify or distort reality.

The chapter therefore proceeds from foundations to design and then to accountable execution planning. Sections 3.1-3.4 define the nature of business research, philosophical assumptions and the conversion of practical concerns into researchable questions. Sections 3.5-3.10 develop literature synthesis and the major families of research design. Sections 3.11-3.14 address sampling, measurement, data collection, management and analysis planning. Sections 3.15-3.18 develop ethics, data protection, generative AI, proposals, project management and transparent dissemination. Sections 3.19-3.20 apply the framework to African and comparative international cases. The final sections integrate Christian moral reasoning, provide a professional research-design toolkit, identify recurring errors, and hand the reader to Chapter 4, where sound designs meet statistical evidence.

Opening Case: Nigeria's YouWiN! Competition and the Difference Between Selection and Causation

In 2011, Nigeria launched the Youth Enterprise With Innovation in Nigeria programme, commonly known as YouWiN!, a national business-plan competition intended to identify promising entrepreneurs and support business creation and growth. Thousands of applicants entered, and winners could receive grants that were commercially large relative to the scale of most early-stage firms. The programme created an attractive policy story: find talented entrepreneurs, finance them, and observe whether they grow. Yet the central research question was more difficult than the policy narrative. If funded entrepreneurs later employed more workers than unfunded entrepreneurs, did the grants cause the difference, or were winners simply more capable from the beginning?

David McKenzie's evaluation turned this practical policy question into a research-design problem. Among applicants who reached a relevant stage of the competition, a portion of grants was randomly assigned, creating treated and comparison groups that were similar in expectation before the intervention. The study tracked applicants over five years and reported that winning increased firm entry and survival, sales, profits and employment, including an increase of more than 20 percentage points in the probability of operating a firm with ten or more workers (McKenzie, 2017). The study therefore did more than describe successful winners. Random assignment created a counterfactual basis for estimating what would plausibly have happened to comparable applicants without the grant.

The methodological lesson is not that randomized experiments are always the best business design. Randomization was feasible because scarce grants had to be allocated among eligible applicants. Many questions cannot ethically or practically be randomized. A researcher cannot randomly assign countries to wars, firms to corrupt legal systems, workers to severe discrimination, or economies to currency crises. Even in YouWiN!, the experimental estimate concerns applicants and the intervention as implemented in that context. It does not prove that the same grant would generate the same effect for every entrepreneur, sector, country or economic period.

The case also reveals the chain of decisions hidden behind a headline result. Who counted as the population of interest: all Nigerian youth, all entrepreneurs, all applicants, or only applicants at a certain competition stage? How were outcomes such as employment, profits and survival measured? How much attrition occurred over five years? Could respondents misreport performance because they knew their treatment status? What ethical obligations arose when some eligible applicants received resources and others did not? Which parts of the intervention were inseparable from the competition's screening and support structure? How should a policymaker distinguish the study's causal estimate from a promise that scaling the programme will reproduce the same effect?

> **EVIDENCE CHECK | A strong finding begins with a strong comparison**
> Observing that funded firms grew does not by itself show that funding caused growth. A causal claim requires a credible account of what would have happened without the intervention. Randomization is one powerful way to construct that comparison, but causal credibility can also be strengthened through well-designed quasi-experiments, longitudinal designs, careful matching, natural experiments and triangulation. The method must fit the question and the feasible ethical conditions.

The opening case gives this chapter its central problem: **How can a researcher design evidence that is fit for the claim, transparent about uncertainty and morally responsible toward the people and institutions from whom the evidence is obtained?**

3.1 Why Research Methods Matter in Business and Trade

Chapter 2 established that professional communication should distinguish claims from evidence, adapt to audiences without distorting material facts, and represent uncertainty honestly. The next question is prior to communication: where does defensible evidence come from? Research methods answer that question by specifying how observations are generated, selected, measured, compared, analyzed and interpreted. Good communication cannot rescue weak design. A polished report based on a biased sample, invalid measure or unsupported causal inference remains weak evidence.

Business research matters because managers and policymakers routinely act under uncertainty. A retailer asks why customer retention fell. An exporter asks whether certification cost or exchange-rate volatility is constraining market entry. A bank asks whether a new credit-scoring process unfairly excludes viable borrowers. A ministry asks whether an entrepreneurship programme creates additional jobs or merely subsidizes firms that would have grown anyway. An investor asks whether a reported association between governance quality and return survives alternative explanations. Each question contains a practical decision, but research is needed when the answer is not already known and when a systematic inquiry can reduce uncertainty.

The approved scope of this chapter is research design. It teaches how to formulate questions, connect them with evidence, choose designs, sample, measure, collect data, plan analysis, protect participants and communicate methodological limits. It deliberately does not reteach the descriptive and inferential statistical techniques owned by Chapter 4, the predictive and prescriptive analytics owned by Chapter 5, or the actual capstone execution process owned by Chapter 22. A Chapter 3 proposal may state that a later study will estimate a multiple-regression model or conduct thematic analysis, but the statistical mechanics of estimation and inference belong to the relevant specialist chapter.

Research methods are also a professional discipline of intellectual humility. A method does not make a conclusion true. Methods organize exposure to error. They help researchers ask where a conclusion could be wrong, what comparison is relevant, what information is missing, what assumptions are required, and what alternative explanations remain. Rigour therefore means disciplined vulnerability to correction rather than ceremonial use of technical vocabulary.

3.1.1 Research, inquiry, evaluation and analytics

The term *research* is used broadly. For this chapter, research is **systematic, transparent and ethically accountable inquiry designed to produce or test knowledge claims using evidence**. Systematic does not mean rigid. It means that choices are reasoned and connected. Transparent means that a competent reader can understand how the evidence was generated and how the conclusion follows. Accountable means that researchers remain responsible for accuracy, participant protection and the consequences of their conduct.

Research overlaps with, but is not identical to, several professional activities. Market intelligence may emphasize timely decision support rather than generalizable scholarship. Programme evaluation asks how an intervention performs against stated objectives and under what conditions. Audit tests evidence against standards and controls. Analytics extracts patterns or predictions from data. Consulting diagnoses a client's problem and recommends action. These activities can use research methods, and some can themselves generate publishable research, but their purposes, audiences and standards of independence may differ.

Table 3.1. *Research purposes, typical questions and characteristic methodological risks.*

Activity	Primary purpose	Typical question	Main methodological risk
Academic research	Develop, test or refine knowledge	What relationship, mechanism or meaning can be supported beyond this immediate decision?	Methodological ritual detached from practical significance
Applied business research	Reduce uncertainty for a defined organizational decision	Why is retention falling, and which explanation is most credible?	Sponsor pressure, narrow framing, inaccessible data
Programme or policy evaluation	Assess implementation, outcomes, impact or value	Did the programme change outcomes relative to a credible counterfactual?	Confusing activity counts with outcomes or causal impact
Market intelligence	Support timely competitive or market decisions	What is changing in customer demand, competitors or regulation?	Speed displacing verification and representativeness
Business analytics	Describe, predict or optimize using data	Which customers are at risk of churn?	Treating model accuracy as causal explanation or moral legitimacy
Audit/compliance review	Test evidence against criteria and controls	Was the process compliant and are reported records reliable?	Checklist compliance without substantive risk analysis

Source/Note: Author's synthesis from sources cited in the surrounding text.

MANAGERIAL INSIGHT | Ask what kind of answer the decision actually needs
A board deciding whether to enter a market may need credible descriptive estimates, scenario ranges and institutional analysis rather than a causal experiment. A government deciding whether a subsidy created jobs needs a counterfactual design. A founder exploring why users abandon an application may first need qualitative discovery. Methodological sophistication is not the number of techniques used. It is the fit between the question, claim and evidence.

3.1.2 Basic, applied and engaged research

Basic research seeks general understanding without requiring an immediate operational decision. Applied research is directed toward a practical problem. Business and trade scholarship often combines the two. A study of how institutional uncertainty affects exporter investment can contribute to theory while helping firms and policymakers. Engaged research further involves practitioners, communities or organizations in problem definition, data generation or interpretation. Engagement can improve relevance and contextual understanding, but it can also create conflicts of interest or pressure to suppress inconvenient results. The design must therefore make roles and decision rights explicit.

3.1.3 The research cycle is iterative

Textbooks often depict research as a straight line: select a topic, review literature, choose a method, collect data, analyze and write. Real inquiry is more recursive. A literature review may reveal that the proposed question has already been answered. A pilot interview may expose ambiguous constructs. An ethics review may show that recruitment would create coercive pressure. A secondary dataset may not contain a valid indicator. An initial analysis may reveal that a planned model cannot support the intended inference. Iteration is legitimate when it improves alignment and is documented. Undisclosed post-hoc changes that make results look stronger are not.

3.2 Philosophical Foundations: What Exists, What Can Be Known and How Claims Are Justified

Methodological choices rest on assumptions about reality and knowledge even when researchers do not name them. Philosophy matters because a study can become internally incoherent if it asks an interpretive question while demanding only context-free measurement, or claims universal causal certainty from a design that can support only local association. Philosophical literacy helps the researcher understand what a method assumes and what kind of claim it can reasonably support.

3.2.1 Ontology, epistemology, axiology and methodology

Ontology concerns assumptions about the nature of reality. Does the phenomenon exist independently of how people interpret it, or is its social meaning partly constituted through interpretation? **Epistemology** concerns what counts as knowledge and how claims can be justified. **Axiology** concerns the role of values in inquiry, including which problems are considered important and how the researcher handles moral commitments. **Methodology** is the reasoned logic connecting these assumptions to a strategy of inquiry. Methods are the concrete techniques used within that strategy.

The distinction prevents a common error: naming software or a data-collection instrument as a methodology. “We used SPSS” is not a methodology. “We used a cross-sectional survey design to estimate associations among operationalized constructs, with stratified sampling and pre-specified regression analysis” describes a research design and methods. The philosophical justification explains why those choices fit the knowledge claim.

3.2.2 Major research orientations

No short table can capture the diversity within philosophical traditions, but a disciplined comparison is useful.

Table 3.2. *Comparing major research orientations in applied business inquiry.*

Orientation	Ontological emphasis	Epistemic emphasis	Typical methodological implications	Important caution
Postpositivist	A reality exists independently of the researcher, but observation is fallible and theory-laden	Hypotheses should face empirical tests, measurement error and alternative explanations	Structured measurement, comparison, falsifiable propositions, statistical uncertainty, replication	Does not justify pretending that measures are perfect or that all relevant phenomena are readily quantifiable
Interpretive / constructivist	Social reality includes meanings created and negotiated by actors	Understanding requires interpretation of situated meanings and practices	Interviews, observation, texts, case analysis, reflexivity, contextual explanation	“Multiple meanings” does not mean every interpretation is equally well supported

Orientation	Ontological emphasis	Epistemic emphasis	Typical methodological implications	Important caution
Critical realist	Real structures and mechanisms exist, while observations are partial and events depend on context	Inquiry seeks causal mechanisms and conditions rather than only regularities	Case comparison, mixed evidence, process tracing, retroduction, attention to open systems	Mechanisms can become speculative if they are not tied to discriminating evidence
Pragmatic	Research choices are judged by how well they address the inquiry and consequences of knowledge use	Methods can be combined when their assumptions are made coherent enough for the problem	Mixed methods, design flexibility, problem-centred inquiry	Pragmatism is not "anything goes"; convenience is not a methodological justification
Critical / emancipatory	Social reality is shaped partly by power, structures and historical inequalities	Knowledge production should examine domination, voice and interests	Participatory, critical discourse, action research, reflexive and power-conscious designs	Advocacy can outrun evidence if desired conclusions are insulated from challenge

Source/Note: Author's synthesis from sources cited in the surrounding text.

These orientations are not political parties that researchers must join permanently. Nor should paradigms be treated as personality types. Guba and Lincoln (1994), Crotty (1998), Burrell and Morgan (1979), Morgan (2007) and later methodologists show that philosophical categories are useful precisely because they expose assumptions and tensions. A researcher should state assumptions only to the depth needed for the study and should avoid ceremonial claims such as "this study adopts positivism because it uses a questionnaire." Questionnaires can be used within different epistemic projects.

> **CRITICAL DEBATE | Must a study declare one paradigm?**
> Some traditions expect an explicit paradigm because assumptions about reality, knowledge and values affect design. Others argue that rigid paradigm labels can obscure the actual logic of inquiry. The defensible position is neither compulsory label-collecting nor philosophical silence. Researchers should state the assumptions that materially affect the question, evidence and inference, and should show how those assumptions are reflected in the design.

3.2.3 Theory and evidence

A theory is a structured explanation of relationships among concepts that helps explain or predict phenomena within stated boundaries. Evidence consists of observations or records used to evaluate claims. Theory helps researchers decide what to observe and why; evidence can support, weaken, refine or complicate theory. Neither should dominate absolutely. Data without conceptual structure can produce disconnected patterns, while theory protected from inconvenient evidence becomes dogma.

Deduction moves from theory toward testable expectations. Induction develops patterns or concepts from observations. Abduction moves iteratively between surprising observations and plausible explanations, asking what explanation would make the observed pattern intelligible. These modes can coexist in one project. A case researcher may begin with sensitizing concepts, discover an unexpected pattern and then compare competing explanations. A quantitative researcher may test a deductive model and use residual patterns to generate new hypotheses.

3.2.4 Positive and normative claims

Business and trade research often mixes empirical and normative questions. "Firms with export experience are more likely to survive entry into a new market" is empirical. "Government should subsidize exporters" is normative and requires values plus evidence about effects, costs and alternatives. "A policy is legal" is a legal claim, not proof that it is effective or morally legitimate. Research can inform moral reasoning but cannot derive moral obligation from observed frequency alone. This distinction will remain important in the chapter's ethics section.

3.3 From a Practical Problem to a Researchable Problem

A research project usually begins with discomfort: sales are falling, a programme appears ineffective, a corridor is congested, employees are leaving, an exporter cannot secure finance, or policy outcomes differ across regions. The practical problem is the situation that matters to a stakeholder. The research problem is the specific uncertainty or knowledge deficit that must be investigated to understand or respond to that situation.

3.3.1 Practical problem, research problem and decision problem

Suppose a Rwandan food exporter experiences declining repeat orders from a regional market. "Repeat orders are declining" is a practical problem. Several research problems could explain it: quality inconsistency, landed-cost increases,

channel-service failures, changes in buyer preferences, customs delays or competitor actions. “Should we withdraw from the market?” is a decision problem. A defensible study narrows the research problem before recommending action.

A strong research problem usually includes four elements: the observed or credible condition; the specific uncertainty; why the uncertainty matters; and the population, setting or process to which the inquiry applies. It should not contain the answer in advance. “The problem is poor leadership causing employee turnover” is biased if causation has not been established. Better: “Turnover among technical staff has risen, but the relative contribution of supervisory practices, compensation, workload and external labour-market opportunities is not established.”

3.3.2 Problem statements

A problem statement should be evidence-led. It establishes what is known, what remains uncertain, and why the uncertainty is consequential. It should avoid inflated claims such as “no study has ever examined...” unless a systematic search can support them. It should also avoid replacing evidence with moral urgency. A topic may be morally important and still require precise research design.

> **PRACTICE TOOL | The four-sentence problem test**
> Sentence 1 establishes the consequential situation with credible evidence. Sentence 2 identifies what existing knowledge explains. Sentence 3 identifies the unresolved uncertainty, contradiction or contextual limitation. Sentence 4 states why resolving that uncertainty matters for theory, policy or professional decision-making. If the four sentences cannot be written without vague nouns such as “issues,” “challenges” or “impact,” the problem is probably not yet sufficiently specified.

3.3.3 Purpose statements and objectives

The **purpose statement** tells the reader what the study will do and in what setting. Objectives decompose that purpose into specific analytical tasks. Good objectives use verbs that correspond to evidence: *describe, estimate, compare, test, explain, explore, interpret, evaluate, develop, validate*. Vague verbs such as *understand* can be useful in qualitative inquiry but should be made more specific about what understanding will concern.

Objectives should be collectively sufficient but not redundant. If the purpose is to examine drivers of SME export persistence, one objective might describe persistence patterns, another estimate associations with firm capabilities and market conditions, and a third explore how exporters interpret regulatory uncertainty. This would imply a mixed or multi-source design. It would not justify an objective to “recommend policy” unless the evidence base includes a reasoned bridge from findings to feasible options.

3.3.4 Research questions

A good research question is specific enough to guide design but open enough to require evidence. It identifies the phenomenon, population or unit, relationship or meaning, and context where these matter. Descriptive questions ask what exists or how much. Associational questions ask how variables covary. Causal questions ask whether an exposure or intervention changes an outcome. Process questions ask how change occurs. Interpretive questions ask how participants understand or experience a phenomenon. Comparative questions ask how patterns differ across cases or conditions.

Table 3.3. *From weak questions to researchable questions.*

Weak question	Why it is weak	Stronger formulation	Likely design implication
What is the impact of digitalization on SMEs?	“Impact,” “digitalization” and “SMEs” are undefined and the causal claim is premature	How is adoption of cloud-based inventory software associated with stockout frequency among formal retail SMEs in Kigali during 2026?	Quantitative observational design, possibly longitudinal if causal direction matters
Why is cross-border trade difficult?	Overbroad, normative and context-free	How do small-scale traders on a specified corridor describe the administrative, payment and transport frictions that most constrain transaction completion?	Qualitative interviews/observation with purposive sampling
Does training work?	Intervention, outcome and population unspecified	What is the effect of a six-month export-readiness programme on first export entry among eligible firms compared with eligible non-participants?	Experimental or quasi-experimental evaluation if feasible
What strategies make firms successful?	Success and strategy undefined, high reverse-causality risk	Which pre-entry capabilities distinguish firms that persist in a target export market for at least three years from firms that exit within the first year?	Longitudinal/comparative design with careful selection controls

Source/Note: Author’s synthesis from sources cited in the surrounding text.

3.3.5 Hypotheses

A hypothesis is a testable expectation about a relationship, difference or parameter. Hypotheses are appropriate when theory and prior evidence justify sufficiently specific expectations and the design can measure the relevant constructs. They are not mandatory in exploratory qualitative research and should not be invented merely because a template requires them.

A directional substantive hypothesis might state: *H1: Higher supplier delivery reliability is associated with lower production interruption frequency, controlling for firm size and inventory buffer.* A null hypothesis used in statistical testing would represent a specific no-effect or no-difference proposition. Chapter 4 explains hypothesis testing and why researchers do not "prove" or "accept" a null hypothesis through non-significance alone.

3.3.6 The alignment matrix

Methodological quality is often visible in alignment. Each question should have evidence capable of answering it, and every collected variable or interview domain should serve a research purpose.

Table 3.4. *Research alignment matrix linking questions to evidence and validity threats.*

Research question	Construct / phenomenon	Evidence source	Sampling unit	Analysis plan	Main validity threat
How frequently do formal SMEs report customs delays above five days?	Clearance delay	Administrative records plus firm survey	Shipment / firm	Descriptive estimates with defined denominator	Missing records and recall error
Which firm characteristics are associated with delay exposure?	Firm and shipment characteristics	Survey + administrative linkage	Firm / shipment	Multivariable association model	Confounding and measurement error
How do exporters explain the commercial consequences of unpredictable clearance?	Experienced mechanisms and adaptation	Semi-structured interviews	Export manager	Reflexive thematic analysis	Selection and researcher interpretation

Source/Note: Author's synthesis from sources cited in the surrounding text.

Figure 3.2. The research-alignment spine from problem to defensible inference.
Source: Author's synthesis.

> **MANAGERIAL INSIGHT | Alignment is more valuable than methodological decoration**
> A modest study with one clear question, a defensible sample and a transparent analysis can be more useful than a project that collects surveys, interviews and secondary data without explaining how they answer the same problem. Every additional method creates cost, error opportunities and ethical responsibilities. Add evidence streams because they improve inference, not because methodological variety looks sophisticated.

3.4 Literature Searching, Evidence Synthesis, Frameworks and Research Gaps

A literature review is not a ceremonial chapter written after the research question is fixed. It is part of problem definition and design. It shows what is already known, which concepts have been used, how prior studies generated evidence, where findings converge or conflict, and which uncertainty genuinely remains. In business fields where scholarship is fragmented across management, economics, information systems, law and policy, the quality of the search strategy materially affects the quality of the claimed gap.

3.4.1 Search before claiming novelty

A defensible search begins with concepts rather than a single preferred phrase. Researchers identify synonyms, related terms, controlled vocabulary where available, and spelling variants. Boolean logic is then used deliberately: **AND** narrows by requiring concepts, **OR** broadens within a concept family, and **NOT** should be used cautiously because it can remove relevant records. Phrase searching, field searching, truncation and date or document-type filters can improve precision when used transparently.

A business researcher studying digital payments and informal trade might combine terms such as *(mobile money OR digital payment* OR fintech) AND (informal trader *OR microenterprise* OR small-scale cross-border trader*) AND (Africa OR specific countries/regions)*. Searching only the preferred final title risks missing literatures that use different terminology.

Strong searches triangulate sources. Bibliographic databases provide peer-reviewed literature; official statistical and legal sources provide primary evidence; working-paper repositories can reveal emerging research; citation chasing

identifies influential and related studies; and targeted grey-literature searching can locate policy and industry evidence. Search engines are useful for discovery but do not by themselves define a reproducible scholarly search.

3.4.2 Backward and forward citation searching

Backward citation searching examines a relevant study's references. Forward citation searching identifies later works that cited it. Together they reveal intellectual lineages, replications, critiques and conceptual developments that keyword searches can miss. Citation counts can signal influence but should not be confused with quality or truth. Highly cited work can be methodologically weak, controversial or influential precisely because it is disputed.

3.4.3 Evidence synthesis, not source accumulation

Synthesis organizes studies around claims, mechanisms, methods, contexts and contradictions. A strong paragraph may show that several studies find an association, but that the association weakens in longitudinal designs; or that survey research identifies a pattern while qualitative work explains a plausible mechanism. The researcher should ask *why* findings differ. Differences may reflect samples, measurement, time periods, institutional conditions, estimation strategies or genuinely heterogeneous effects.

Snyder (2019) emphasizes that literature reviews can themselves be research methods when conducted with a transparent approach. Tranfield, Denyer, and Smart (2003) argued for more systematic evidence synthesis in management while recognizing that management questions often require context-sensitive interpretation. Review methodology should therefore fit the review question rather than mechanically copying a biomedical template.

3.4.4 Narrative, integrative, scoping and systematic reviews

A **narrative review** provides an interpretive synthesis and can be highly valuable when the purpose is conceptual development or historical integration, but it must still disclose how literature was selected sufficiently to avoid invisible cherry-picking. An **integrative review** intentionally combines diverse theoretical and empirical evidence to generate new frameworks or perspectives. A **scoping review** maps the extent, characteristics and conceptual boundaries of a field, often when questions are broad and evidence heterogeneous. A **systematic review** uses a pre-specified and transparent process to identify, screen, appraise and synthesize evidence for a focused question.

PRISMA 2020 provides an internationally recognized reporting guideline for systematic reviews (Page et al., 2021). It improves transparency about why a review was done, what was searched, what was included and what was found. PRISMA should not be used as a quality badge detached from the underlying review design. A flow diagram cannot compensate for a weak search, inappropriate inclusion criteria or uncritical synthesis.

> **EVIDENCE CHECK | Review labels do not create rigour**
> A document called a “systematic review” can still be biased. A narrative review can still be intellectually rigorous. The relevant questions are whether the search and selection logic fit the purpose, whether exclusions are defensible, whether evidence quality is examined, whether synthesis respects methodological differences, and whether readers can understand how the conclusion was produced.

3.4.5 Eligibility, critical appraisal and synthesis planning

Criteria should be derived from the question. Typical dimensions include population or unit, phenomenon/exposure/intervention, comparator, outcomes, study design, geography, time period, language and publication type. Criteria should not be written after screening in order to remove inconvenient evidence. Any post-protocol change should be documented and justified.

Eligibility is only the first quality gate. After potentially relevant studies have been identified, the researcher should critically appraise whether each design can support the claim for which it is being used. For intervention studies, this may require attention to allocation, implementation fidelity, attrition, outcome measurement and selective reporting. For observational studies, confounding, selection, measurement and temporal ordering are central. Qualitative studies require appraisal of sampling logic, contextual adequacy, reflexivity, analytic transparency and the connection between interpretation and evidence. A universal numerical score is rarely a substitute for design-specific judgment.

Data extraction should also be planned before synthesis. A transparent extraction matrix records study context, sample, constructs, measures, design, analytic approach, main findings, uncertainty, limitations and relevance to the review question. When a review has more than one reviewer, independent screening or verification can reduce unnoticed selection errors, but duplicate review is not a ritual requirement for every low-stakes assignment. What matters is an auditable process proportionate to the claim.

Synthesis should preserve meaningful differences among studies rather than average away incompatible questions. Meta-analysis is appropriate only when effect estimates, populations, interventions or exposures, and inferential targets are sufficiently comparable; otherwise a structured narrative, tabular or configurative synthesis may be more defensible.

Researchers should also consider publication and availability biases: null, adverse or commercially inconvenient findings may be less visible than positive results. Review credibility therefore depends on transparent search, appraisal, extraction and synthesis decisions, not on the review label alone (Page et al., 2021; Snyder, 2019; Tranfield et al., 2003).

3.4.6 Theoretical and conceptual frameworks

A **theoretical framework** uses an established theory or coherent set of theoretical propositions to organize constructs and expected relationships. A **conceptual framework** is the researcher's reasoned representation of the concepts and relationships most relevant to the study, often integrating multiple sources. The two overlap in practice, but neither should be a decorative diagram.

A useful framework performs work. It identifies what should be observed, clarifies relationships, guides measure selection or interview domains, and provides a structure for interpreting findings. A framework becomes dangerous when it prevents the researcher from noticing evidence that does not fit. Qualitative designs may therefore use “sensitizing concepts” more lightly than hypothesis-driven studies.

3.4.7 What counts as a research gap?

A gap is not simply an empty geographic cell. “No study has examined this topic in Country X” can justify research only if the context plausibly changes theory, mechanism, measurement or decision relevance. More defensible gaps include a contradictory evidence gap, a methodological gap, a measurement gap, a mechanism gap, a population gap with theoretical significance, a temporal or institutional-change gap, or an implementation gap between known efficacy and uncertain real-world performance.

Table 3.5. *A practical taxonomy of defensible research gaps.*

Gap type	Legitimate question	Weak version to avoid
Contradictory evidence	Why do credible studies reach different estimates under different designs or contexts?	“Studies disagree, therefore more research is needed.”
Methodological	Does a relationship observed cross-sectionally persist under longitudinal or quasi-experimental design?	“Previous studies used questionnaires; I will also use a questionnaire.”
Mechanism	Through what process does an observed association arise?	Adding interviews without specifying what mechanism they will examine
Measurement	Are existing indicators valid and comparable in this institutional context?	Renaming an established scale without validation
Contextual boundary	Does a theory's mechanism change under a materially different institutional or market setting?	Country novelty alone
Implementation	Can an intervention shown effective under controlled conditions work at operational scale?	Assuming efficacy equals scalability

Source/Note: Author’s synthesis from sources cited in the surrounding text.

3.5 Research Design as the Architecture of Inference

A research design is the logic that connects a question to evidence and a conclusion. It specifies what will be observed, from whom or what, when, under which comparison, through which measures, and with what analysis. Classic work on experimental and quasi-experimental design emphasizes that threats to inference are problems of research architecture, not only statistical technique (Campbell & Stanley, 1963; Shadish et al., 2002). Design therefore precedes software. Statistical sophistication cannot reconstruct a comparison that the data-generating process never created.

3.5.1 Unit of analysis and level of inference

The **unit of analysis** is the entity about which the study makes claims: person, team, firm, transaction, shipment, market, country, policy or network. The unit of observation is the level at which data are recorded. They can differ. A study may observe individual employee responses but infer about teams; it must then account for clustering and justify aggregation. A country-level association between corruption indices and trade cannot be used to infer how a particular manager behaves. Such cross-level errors are ecological or atomistic fallacies depending on the direction of inference.

3.5.2 Time horizon

A **cross-sectional** design observes a phenomenon at one time or over a short window. It is efficient for describing prevalence and associations but often weak for temporal ordering. A **longitudinal** design observes units over time and can study change, sequence and persistence. Panel designs follow the same units; repeated cross-sections sample the population at different times; cohort designs follow groups defined by a shared starting condition. Longitudinal data

improve temporal evidence but do not automatically establish causality because confounding and selective attrition can remain.

3.5.3 Observational and intervention designs

Observational designs study exposures and outcomes without assigning the exposure. Intervention designs deliberately assign or introduce a condition. The distinction matters because treatment assignment affects causal credibility. Random assignment can balance observed and unobserved characteristics in expectation, while observational studies must rely more heavily on design, theory and adjustment to address confounding.

QUESTION EMPHASIS	DESIGN FAMILIES TO CONSIDER
How much / how often / what pattern?	Survey, census, administrative/secondary data, descriptive longitudinal design
Does X change Y?	Randomized experiment or a credible quasi-experimental/longitudinal counterfactual design
How / why / what does it mean in context?	Case study, interviews, ethnography, qualitative observation, process-focused design
Do we need magnitude and mechanism together?	Mixed methods when integration is necessary and explicitly planned

Figure 3.3. Question-first research-design selection.
Source: Author's synthesis.

3.6 Quantitative Research Designs

Quantitative research represents phenomena using variables and numerical observations so that magnitudes, differences, associations or effects can be estimated. The defining feature is not the presence of a spreadsheet. It is a measurement and inference logic in which concepts are operationalized and analyzed numerically.

3.6.1 Descriptive survey designs

Surveys can estimate characteristics, attitudes, practices or constraints for a defined population when the sampling frame, response process and measurement are defensible. The World Bank Enterprise Surveys illustrate a rigorous business-survey architecture: formal surveys generally use stratified random sampling by firm size, sector and geographic region, with explicit eligibility criteria and sampling weights. The methodology also distinguishes the universe of inference from firms not covered by the design, including many informal firms unless a separate informal survey is conducted.

The practical lesson is that "we surveyed 300 firms" is methodologically incomplete. Readers need to know which firms could have been selected, how they were selected, who responded, how nonresponse was handled, whether weights are needed, and which claims the sample can support.

3.6.2 Correlational and explanatory observational designs

Observational studies can examine associations among variables and test theoretically specified relationships. They are central to business research because many exposures cannot be assigned. Yet association is not causation. If firms using sophisticated analytics are more productive, analytics may improve productivity, productive firms may be more able to invest in analytics, or a third factor such as management quality may cause both.

Design can strengthen inference before analysis begins. Researchers can measure plausible confounders, establish temporal ordering, use repeated observations, select comparison groups, exploit exogenous policy changes, or define a clear causal model. Chapter 4 later develops correlation, regression and statistical inference; Chapter 3's concern is whether the design creates interpretable comparisons.

3.6.3 Randomized experiments

In a randomized controlled experiment, eligible units are assigned by a random process to treatment conditions. Randomization supports causal inference because, with adequate implementation and sample size, treatment groups are comparable in expectation before treatment. Researchers must still guard against noncompliance, spillovers, attrition, treatment contamination, outcome measurement bias and selective reporting.

Randomization can occur at the individual, firm, branch, community or other cluster level. Cluster randomization may be necessary when an intervention is delivered collectively, but it reduces effective information because units within clusters are correlated. Sample-size planning must account for the randomization level and anticipated effect, not simply count individual observations.

3.6.4 Field, laboratory and online experiments

Laboratory experiments offer control and precise manipulation but may simplify organizational context. Field experiments embed intervention in real settings and often improve realism while sacrificing some control. Online experiments, including A/B tests, can be operationally powerful but raise ethical questions when users do not reasonably expect experimentation or when high-stakes outcomes are manipulated. A commercial platform's ability to randomize does not eliminate obligations of informed governance, risk assessment and privacy.

The choice among laboratory, field and online settings therefore requires an explicit trade-off between control, realism, implementation fidelity and participant risk. A tightly controlled laboratory effect may fail when organizational routines, incentives or competing demands are restored. A field experiment may reveal operationally meaningful effects while introducing spillovers, treatment variation and measurement noise. Online experimentation can achieve large samples quickly, but scale magnifies even small harms and can make repeated testing invisible to users. The researcher should state what feature of the setting strengthens the intended inference and what new threat the setting introduces.

3.6.5 Quasi-experiments and natural experiments

Quasi-experimental designs estimate causal effects without researcher-controlled randomization by exploiting comparison structures such as interrupted time series, regression discontinuity, difference-in-differences, matched comparisons or instrumental variables. These designs can be powerful when their assumptions are credible, but each substitutes assumptions for the balance guaranteed in expectation by randomization. The researcher must explain why the comparison approximates the counterfactual and what would falsify that logic.

A **natural experiment** is not merely an event that happened naturally. It refers to an external process that generates treatment variation plausibly unrelated to potential outcomes in a way useful for causal inference. A policy boundary, lottery, threshold or sudden institutional change can sometimes provide such variation. Researchers should be cautious about using the label when assignment is clearly strategic or self-selected.

> **CRITICAL DEBATE | The hierarchy of methods is question-dependent**
> Randomized experiments have exceptional strength for many causal questions, but they are not universally superior. A randomized trial can answer a narrow intervention question while missing historical meaning, implementation mechanisms or rare harms. A well-executed qualitative case can be superior for explaining process. A census can be superior to a sample for describing a small finite population. Methodological maturity means matching design strength to the claim rather than treating one method as a universal status hierarchy.

3.7 Qualitative Research Designs

Qualitative research investigates meaning, process, experience, practice, context and mechanism using evidence that is not reduced primarily to numerical variables. It is not “research without numbers,” nor is it inherently subjective in the sense of being unconstrained by evidence. Rigorous qualitative inquiry makes the researcher's interpretive role visible, samples deliberately, documents analytic decisions, tests interpretations against the data, and represents divergent or negative cases rather than selecting only vivid quotations (Maxwell, 2013; Patton, 2015).

3.7.1 Case study research

A case study investigates a bounded contemporary phenomenon in depth and in context, commonly using multiple evidence sources (Yin, 2018). The “case” must be defined. A company, merger, policy implementation, supply-chain disruption, trade corridor, market-entry process or interorganizational network can be a case. A study that merely interviews ten managers at one firm is not automatically a case study unless the firm or process itself is the analytical case and the evidence is organized to understand it.

Case selection follows theoretical rather than merely convenient logic. A **critical case** may test a proposition under conditions where it should be especially likely or unlikely. An **extreme or revelatory case** provides access to an unusual phenomenon. A **typical case** illustrates a common configuration. A **multiple-case design** permits structured comparison and can strengthen theoretical inference when cases are selected for replication logic rather than accumulated indiscriminately (Eisenhardt, 1989; Yin, 2018).

Case studies can include quantitative evidence. Financial records, transaction data, timelines and survey results can be integrated with interviews and documents. The defining feature is the bounded case and contextual explanatory logic, not the exclusivity of a data type.

3.7.2 Ethnography and organizational observation

Ethnographic research seeks to understand practices, meanings and social organization through sustained engagement, observation and participation in a field setting. In business and trade, ethnographic approaches can reveal

how formal policies are actually enacted, how informal exchange operates, how technology is appropriated in daily practice, or how traders navigate institutions that look different in policy documents than at the border.

Ethnography demands unusually careful reflexivity. Access can depend on gatekeepers; the researcher's presence can alter behavior; close relationships can create loyalty conflicts; and fieldnotes may contain sensitive details about persons who never formally sat for an interview. Consent may need to be treated as an ongoing relational process rather than a single signature.

3.7.3 Grounded theory

Grounded theory is a family of approaches intended to develop explanatory categories and theory through iterative data collection and analysis. Classic and constructivist variants differ in epistemological assumptions, but both reject the idea that grounded theory is simply “coding interviews into themes.” Constant comparison, theoretical sampling, memo writing and category development are central. Researchers should not label any inductive thematic study as grounded theory when they do not use its defining logic.

3.7.4 Phenomenological approaches

Phenomenological research examines lived experience and the meaning structures through which a phenomenon is experienced. It may be appropriate for questions such as how displaced entrepreneurs experience rebuilding commercial identity after conflict or how employees experience algorithmic performance management. Phenomenology is less appropriate when the primary question is the prevalence of an attitude or the causal effect of a policy.

Phenomenological traditions differ in how they treat description, interpretation and the researcher’s prior understanding. A rigorous study should therefore identify the specific phenomenological orientation it is using, recruit participants who have direct experience of the phenomenon, gather sufficiently rich first-person accounts, and make the path from experiential evidence to interpretive claims visible. Reflexive practices, careful attention to language and deliberate examination of cases that complicate the emerging account help prevent the researcher from converting a preferred interpretation into an assumed essence.

Phenomenological findings are normally claims about the structure or meaning of experience within a bounded context, not estimates of prevalence. A study can illuminate how entrepreneurs experience regulatory uncertainty without establishing how common that experience is across all entrepreneurs. If prevalence, group differences or causal effects are also required, the design should add evidence suited to those claims rather than stretching phenomenological evidence beyond its proper inferential role.

3.7.5 Narrative inquiry

Narrative inquiry studies stories as data and as forms through which people organize experience over time. Entrepreneurial biographies, organizational founding narratives and accounts of crisis can reveal identity, temporality and sensemaking. Researchers must distinguish between treating stories as transparent reports of fact and treating them as situated constructions with rhetorical and relational functions.

Narrative analysis can examine sequence, turning points, plot, voice, audience, omission and the relationship between personal accounts and organizational or institutional narratives. In business research, this is especially valuable when the research question concerns how founders explain strategic pivots, how organizations construct legitimacy after failure, or how workers interpret major technological change. The analytic unit may be an entire life story, a bounded organizational episode or a corpus of recurring narratives, and the sampling logic should match that unit.

Narrative coherence is not the same as factual verification. People retrospectively organize events into meaningful stories, sometimes compressing uncertainty or attributing intention after outcomes are known. Where factual chronology matters, researchers should triangulate narratives with documents, records or other accounts while preserving the narrative itself as evidence of meaning and identity. This distinction allows narrative research to remain interpretively rich without treating every remembered detail as an independently verified historical fact.

3.7.6 Qualitative description and reflexive thematic analysis

Not every qualitative project needs a grand methodological label. Qualitative description can provide a disciplined account of experiences or processes using a clear sampling and analytic strategy. Reflexive thematic analysis, associated with Braun and Clarke (2006) and subsequent development of their approach, identifies patterned meaning across a dataset through active researcher interpretation. It differs from mechanical code counting and from versions of content analysis that prioritize replicable coding frequencies.

Gioia, Corley, and Hamilton (2013) propose a data structure commonly used in inductive organizational research, moving from informant-centred first-order terms to researcher-generated second-order themes and aggregate

dimensions. The approach can improve transparency, but its visual data structure should not become a ritual detached from genuine theory development.

Table 3.6. *Comparing major qualitative research designs.*

Qualitative design	Primary intellectual task	Typical evidence	Sampling logic	Common misuse
Case study	Explain a bounded phenomenon in context	Interviews, records, observations, documents, quantitative case evidence	Theoretical, critical, typical, extreme or replication logic	Calling any single-site interview study a case study
Ethnography	Understand situated practices and cultural meaning	Participant observation, fieldnotes, artifacts, interviews	Field/site access and purposive engagement	Short visits described as ethnography
Grounded theory	Develop explanatory categories and theory	Iterative interviews/observations/documents	Theoretical sampling	Using the label for generic thematic coding
Phenomenology	Understand lived experience and meaning	In-depth first-person accounts	Participants with direct experience	Using it for broad organizational-description questions
Narrative inquiry	Analyze stories, temporality and identity	Life histories, organizational narratives, documents	Information-rich narrators or narrative corpus	Treating every story as direct factual evidence
Qualitative description / thematic study	Characterize experiences, practices and patterned meanings	Interviews, focus groups, documents	Purposive maximum variation or relevance	Presenting themes without analytic explanation or negative cases

Source/Note: Author's synthesis from sources cited in the surrounding text.

3.7.7 Saturation and information power

Qualitative sample size should not be justified by universal numerical folklore. **Saturation** is used differently across traditions. Hennink, Kaiser, and Marconi (2017), for example, distinguish code saturation from deeper meaning saturation and show that they can occur at different stages. Their empirical sample counts are not universal rules. Malterud, Siersma, and Guassora (2016) propose **information power**: fewer participants may be required when the study aim is narrow, the sample is highly specific, theory is strong, dialogue quality is high and analysis is intensive.

A defensible qualitative sample-size statement therefore explains the diversity required, the information expected from each case, the analytic approach and the stopping logic. "We interviewed 12 because qualitative studies need 12" is not a methodology.

> **PRACTICE TOOL | A defensible qualitative stopping rule**
> Before fieldwork, specify the diversity or cases that must be represented. During analysis, track whether new interviews add substantively new categories, mechanisms, exceptions or contextual conditions. Stop because the dataset is sufficiently informative for the bounded question, not because an arbitrary number has been reached. Document what "sufficiently informative" meant for the chosen qualitative tradition.

3.8 Mixed-Methods Research: Integration, Not Mere Multiplicity

Mixed-methods research intentionally integrates qualitative and quantitative approaches so that their combination addresses a research problem more effectively than either could alone. Two separate datasets placed side by side do not automatically constitute mixed methods. Integration must occur in design, sampling, data collection, analysis, interpretation or some combination of these stages (Fetters et al., 2013).

3.8.1 Why mix methods?

Methods may be mixed for **complementarity**, when one evidence stream explains or enriches another; **triangulation**, when different approaches assess the same proposition from different vulnerabilities; **development**, when one phase informs the next; **expansion**, when different methods address different dimensions of a complex problem; or **initiation**, when contradiction between methods generates deeper inquiry.

Mixing should not be justified as "qualitative gives depth and quantitative gives breadth" without explaining why both are necessary for the actual research question. Qualitative research can sometimes be broad across cases, and quantitative research can sometimes investigate mechanisms. The justification must concern the information each component contributes.

3.8.2 Core mixed-methods designs

Table 3.7. Core mixed-methods designs and their integration risks.

Design	Sequence and priority	Integration logic	Suitable example	Main risk
Convergent	Qualitative and quantitative components conducted in parallel	Compare, merge or explain convergence/divergence	Survey exporters about border delays while interviewing a purposive subsample about coping mechanisms	Forcing agreement and ignoring contradictory findings
Explanatory sequential	Quantitative first, qualitative second	Follow-up qualitative inquiry explains quantitative patterns	Analyze churn patterns, then interview cases with unexpected trajectories	Second phase becomes anecdotal decoration rather than targeted explanation
Exploratory sequential	Qualitative first, quantitative second	Qualitative findings develop constructs, measures or hypotheses for testing	Explore informal credit practices, develop context-valid items, then survey a larger sample	Prematurely turning nuanced concepts into weak measures
Embedded	One method is nested within a dominant design	Secondary method answers a specific process or implementation question	RCT with interviews on implementation fidelity and participant experience	Embedded component is under-resourced or analytically disconnected
Multiphase	Multiple linked studies across time	Iterative integration across a programme of research	Diagnostic qualitative phase, survey validation, intervention pilot, evaluation	Scope becomes unmanageable and alignment is lost

Source/Note: Author's synthesis from sources cited in the surrounding text.

3.8.3 Integration techniques

Integration can occur through **connecting**, where one phase's sample informs the next; **building**, where findings from one method shape instruments or protocols; **merging**, where datasets are compared or combined; and **embedding**, where one dataset is nested within another. Joint displays can place quantitative results alongside qualitative explanations, but a joint display is useful only when it makes an integrative inference visible.

A mixed-method conclusion should identify what is learned from the combination. If the quantitative component finds that firms with diversified logistics routes have lower delay variance and the qualitative component shows that diversification is effective only where managers maintain active broker relationships, the integrated inference concerns a conditional mechanism. The contribution is not simply "both methods found logistics matters."

3.9 Secondary Data, Administrative Records and Documentary Research

Secondary data research uses evidence originally generated for another purpose or by another actor. It can be a primary research design in its own right. Business and trade scholars increasingly work with customs records, tax data, company financial statements, transaction logs, central-bank series, enterprise surveys, satellite data, procurement records, platform data and archived documents. Secondary research can be more representative, less costly and less intrusive than new data collection, but it inherits the data-generating process of the original source.

3.9.1 Begin with the data-generating process

Before analyzing a dataset, ask why each record exists. Customs data reflect declared transactions under legal and administrative rules. Bank transaction data reflect processed transactions, not all economic exchange. Platform logs record actions the platform can observe, not intentions or activity outside the platform. Company records are shaped by operational definitions and incentives. Survey microdata reflect sampling, nonresponse and questionnaire design. Understanding this process is necessary to interpret missingness, measurement and selection.

3.9.2 Administrative data

Administrative data can offer large samples, longitudinal coverage and reduced respondent burden. They can also contain definitions designed for administration rather than research. A tax record may change because filing rules changed. An apparent increase in registered firms may reflect a registration campaign. A customs-code revision may create a break in a trade series. Researchers should document versioning, classification systems, coverage, revisions and linkage rules.

3.9.3 Official statistics and international comparability

International datasets create powerful opportunities for comparison but can encourage false equivalence. Indicators with identical labels may differ in source quality, price bases, survey frames, informality coverage or revision schedules.

Trade data may differ between exporter and importer records because of valuation, timing, transshipment, partner attribution and reporting errors. A researcher should treat discrepancies as methodological information rather than automatically selecting the more convenient figure.

3.9.4 Company, industry and commercial data

Annual reports are authoritative sources for what a company reports under its reporting obligations, but they are not independent evaluations of the company's social impact or strategic success. Commercial databases can add standardized firm or market information, but researchers need to disclose coverage, vendor methods and licensing constraints. Social media and web-scraped data can reveal digital behavior while systematically excluding people or transactions outside the observed platform.

3.9.5 Document analysis

Documents are both sources of information and social artifacts. Policies, contracts, meeting minutes, advertisements, parliamentary records, court judgments, investor presentations and strategy documents can be analyzed for content, framing, change over time or institutional practice. Researchers should examine authorship, audience, purpose, production context, authenticity and what the document cannot reveal.

3.9.6 Review-based dissertations and evidence syntheses

A dissertation based on secondary literature is still research if it has a defensible methodology. It should state the review question, sources searched, search dates, terms, eligibility criteria, screening procedures, appraisal approach, synthesis method and limitations. “Desk research” is not sufficient methodological description. Review-based research can be especially appropriate where direct participant access is unnecessary, ethically sensitive or impractical, but convenience alone does not justify poor evidence selection.

> **AFRICAN CONTEXT | Formal and informal economies require different inference frames**
> Business datasets built from registered firms can be exceptionally useful while still excluding substantial informal activity. The World Bank Enterprise Surveys explicitly define their standard formal-sector universe and use separate methodologies for informal-sector surveys. African business research should therefore resist treating “available firm data” as equivalent to “the economy.” The relevant universe must be named, and exclusion of informal actors should be considered when interpreting policy or market implications.

3.10 Survey, Interview, Focus-Group, Observation and Document Instruments

Research designs become operational through instruments and protocols. Instrument quality determines what the study actually observes, not what the proposal says it intended to observe.

3.10.1 Survey questionnaires

Questionnaire design begins with constructs and decisions, not with attractive templates. Every item should serve a research purpose. Questions should use language the target population can interpret consistently, avoid double-barrelled formulations, reduce unnecessary recall burden, specify time frames where relevant, and provide response options that are mutually understandable and analytically useful.

Leading questions push respondents toward a desired answer. Loaded questions embed contested assumptions. Double-barrelled items ask more than one question in a single response. Ambiguous frequency terms such as “often” can mean different things to different respondents. Long recall periods increase memory error. Sensitive questions may require privacy-enhancing modes and carefully ordered placement.

Example. “How satisfied are you with the low cost and excellent reliability of our cross-border payment service?” is leading and double-barrelled. Better items would separately ask respondents to rate cost and reliability, with response scales that do not presume excellence.

3.10.2 Response scales

Likert-type items present ordered response categories to statements or questions. Researchers should distinguish an individual Likert item from a multi-item scale that combines several indicators of a construct. Scale points should be clearly labelled when feasible, direction should be consistent unless there is a justified reason otherwise, and “not applicable” should not be forced into the same meaning as a neutral midpoint.

The number of response categories involves trade-offs among discrimination, cognitive burden and comparability. There is no universal rule that five or seven points is always optimal. The response process, construct and population matter more than template tradition.

3.10.3 Translation and cross-cultural measurement

Cross-border research often requires translation. Literal translation can preserve words while altering meaning. A robust process may include professional or bilingual translation, back-translation when useful, committee review, cognitive interviewing and pilot testing. More importantly, researchers must ask whether the construct itself functions similarly across contexts. A management-practice item, trust scale or financial-literacy measure may carry different social meaning even after flawless linguistic translation.

Measurement invariance and cross-cultural comparability become statistical questions in more advanced work, but the design responsibility begins here: do not assume that an instrument validated in one population automatically measures the same construct elsewhere.

3.10.4 Pretesting, cognitive interviewing and piloting

Pretesting checks whether questions, instructions and response options work as intended. **Cognitive interviewing** investigates how respondents understand, retrieve information, judge and map their answers to response options. **Piloting** tests a near-final procedure under realistic field conditions, including recruitment, timing, logistics, data capture and failure modes.

Pilot data should not automatically be merged with the final study. If instruments or procedures change materially, combining phases can contaminate comparability. The proposal should state whether pilot observations will be included and under what conditions.

3.10.5 Semi-structured interviews

An interview guide organizes domains rather than scripting every sentence. Questions should invite concrete experience, examples, processes and interpretations. Follow-up probes can ask “What happened next?”, “Can you give an example?”, “What made that difficult?”, or “Was there a case that worked differently?” Interviewers should avoid turning questions into mini-lectures that reveal the preferred answer.

Power matters. An employee interviewed by a researcher introduced by senior management may reasonably doubt confidentiality. A supplier may fear retaliation. A student may perceive that participation affects grades. Recruitment and interview procedures should therefore separate authority relationships where possible and explain boundaries honestly.

3.10.6 Focus groups

Focus groups use group interaction as data. They are useful for exploring shared language, disagreement, social norms and how participants respond to each other's views. They are less appropriate where participants cannot safely disagree or where confidentiality risk is high. Researchers cannot guarantee that participants will keep others' disclosures confidential, and this limitation should be explained during consent.

Group composition affects discussion. Too much hierarchy may silence junior participants; too much homogeneity may suppress alternative experience; too much diversity may make a sensitive topic unsafe. Moderators should manage dominant voices without converting the session into a sequence of individual interviews.

3.10.7 Observation

Observation can reveal discrepancies between formal procedures and actual practice. Structured observation uses predefined categories; unstructured or ethnographic observation records broader field processes. Researchers should define their role as participant or non-participant, what spaces and events are observable, how consent operates, and how fieldnotes distinguish observation from interpretation.

3.10.8 Documentary and digital-trace protocols

Document studies should specify document types, date ranges, repositories, search rules and extraction procedures. Digital-trace research should define the platform, API or scraping method, time window, identifiers, deleted or inaccessible content and terms-of-service constraints. Public availability is not equivalent to ethical permission for every use, especially when individuals can be re-identified or when context collapse could create harm.

Table 3.8. *Evidence sources, strong uses and characteristic threats.*

Evidence source	Strong use	Characteristic threat	Key design control
Survey	Population estimates, standardized comparisons	Sampling error, nonresponse, measurement error, social desirability	Sampling frame, pretesting, weighting, response-process analysis
Interview	Meaning, process, mechanism, decision logic	Recall, impression management, interviewer influence	Purposeful sampling, probing, reflexivity, triangulation
Focus group	Shared norms, contested views, group language	Conformity, dominance, confidentiality limits	Composition, skilled moderation, explicit privacy limits

Evidence source	Strong use	Characteristic threat	Key design control
Observation	Actual practice and process	Observer effects, selective attention	Field protocol, prolonged engagement, reflexive notes
Administrative data	Large-scale longitudinal operational records	Definition changes, missingness, selection into systems	Data-generating-process audit, metadata, sensitivity checks
Documents	Institutional framing, decisions, historical record	Strategic authorship, incomplete archives	Provenance and context analysis, corroboration
Digital traces	High-frequency platform behavior	Platform selection, bots, opaque algorithms, re-identification	Scope definition, provenance, ethical review, robustness checks

Source/Note: Author's synthesis from sources cited in the surrounding text.

3.11 Population, Sampling Frames and Sample-Size Reasoning

Sampling determines which observations connect the study to a broader population or theoretical claim. Classical sampling theory shows why selection probabilities, frames, strata and clusters matter for design-based inference (Cochran, 1977; Kish, 1965). A large sample selected poorly can be less informative than a smaller sample selected through a defensible design.

3.11.1 Target population, accessible population and sampling frame

The **target population** is the population to which the researcher intends to generalize. The **accessible population** is the portion that can realistically be reached. The **sampling frame** is the operational list or mechanism from which units can be selected. These should not be conflated.

If the target population is all formal manufacturing SMEs in Rwanda but the available frame contains only members of one industry association, the accessible population is narrower. No statistical formula repairs the coverage error. Researchers should either obtain a better frame, redefine the target population or state the limitation clearly.

3.11.2 Probability sampling

Probability sampling gives each eligible unit a known non-zero selection probability, enabling design-based estimates of sampling uncertainty when assumptions are met.

Simple random sampling selects units with equal probability from a complete frame. **Systematic sampling** selects every kth unit after a random start and is efficient when the frame has no harmful periodicity. **Stratified sampling** divides the population into meaningful strata and samples within each, improving precision or ensuring representation of smaller groups. **Cluster sampling** samples natural groups such as branches, villages or firms and then observations within them, often reducing field cost but increasing sampling variance. **Multistage sampling** combines stages, which is common in large surveys.

The World Bank Enterprise Surveys provide a concrete business example of stratification by firm size, sector and location. Such stratification is not cosmetic. It reflects heterogeneous business populations and supports estimates for analytically important subgroups when appropriate weights and design features are retained.

3.11.3 Non-probability sampling

Non-probability sampling does not provide known selection probabilities. It is often appropriate for qualitative and hard-to-reach research, but population prevalence claims must be limited.

Purposive sampling selects information-rich units relevant to the research question. Maximum-variation sampling deliberately captures diversity; criterion sampling requires specified characteristics; critical-case sampling selects analytically decisive cases. **Quota sampling** fills predefined categories without random selection within them. **Snowball or chain-referral sampling** recruits through participant networks and can access hidden populations, but network structure affects who is reached. **Convenience sampling** selects easily available units and should rarely be disguised by stronger labels.

3.11.4 Sample-size reasoning for quantitative estimates

Sample size depends on the estimand, desired precision, expected variability, design, analysis and feasible response. A commonly taught formula for estimating a proportion under simple random sampling is:

$$n0 = z^2 p(1-p) / e^2$$

where z is the critical value for the desired confidence level, p is the anticipated proportion and e is the target margin of error under the formula's assumptions. If the population is finite and sampling without replacement, a finite-population correction can be applied:

$$n = n0 / [1 + (n0 - 1)/N]$$

For a 95% confidence level, z = 1.96. If p is unknown, 0.50 maximizes $p(1-p)$ and therefore gives a conservative size for this specific proportion problem. With e = 0.05:

$$n0 = (1.96^2 \times 0.50 \times 0.50) / 0.05^2 = 384.16$$

If the eligible population is N = 10,000, the finite-population-adjusted size is approximately:

$$n = 384.16 / [1 + (383.16/10{,}000)] = 370.0$$

If only 70% of contacted eligible units are expected to complete the survey, the researcher would need to invite about 370/0.70 = 529 units, assuming the response-rate expectation is credible.

Complex sampling changes the amount of independent information in a sample. In a simple equal-size cluster design, a common planning approximation uses the design effect (DEFF), which depends on the average cluster size m and the intracluster correlation rho. When observations within the same cluster resemble one another, the effective sample size is smaller than the raw count suggests.

$$DEFF \approx 1 + (m - 1)\rho \quad \text{and} \quad n_{clustered} \approx n_{SRS} \times DEFF$$

This expression is a planning approximation, not a universal correction. Unequal cluster sizes, stratification, unequal weights and multistage designs can alter variance in more complex ways. Sample-size planning should therefore reflect the actual selection design and primary estimand rather than multiplying every study by an arbitrary design effect (Cochran, 1977; Kish, 1965).

WORKED EXAMPLE | Why 384 is not a universal sample size
The well-known value near 384 comes from one precision formula under simple random sampling, p = .50, a 95% confidence level and a 5-percentage-point margin of error for a proportion in a very large population. It is not a universal dissertation rule. Cluster sampling can require more observations because of design effects. Detecting a small treatment effect can require far more. Estimating a rare outcome can require far more. A census of a population of 120 firms cannot require 384 firms. Sample size must follow the estimand and design.

3.11.5 Power and minimum detectable effects

For hypothesis tests and experiments, **statistical power** is the probability of detecting a specified effect under defined assumptions when that effect exists. Sample-size planning should specify the effect size or minimum detectable effect, significance criterion, desired power, outcome variability, allocation ratio and clustering where relevant. Power calculations are design-specific and belong conceptually before data collection. Chapter 4 will develop statistical error and inference in detail.

The ethically responsible approach is not to choose an effect size only because it produces an affordable sample. The expected effect should be grounded in theory, prior evidence or the smallest effect that would matter for the decision. Underpowered studies can waste participant time and create noisy evidence; excessively large studies can detect trivial effects and consume unnecessary resources.

3.11.6 Qualitative sample adequacy

Qualitative adequacy depends on the study aim, case diversity, information power, analytic depth and evidence quality. Researchers should state what dimensions of variation need coverage and how they will determine that further data add limited substantive insight. Saturation should be defined rather than invoked as a magic word.

3.11.7 Sampling bias and nonresponse

Coverage bias occurs when parts of the target population are absent from the frame. **Selection bias** occurs when inclusion relates to the variables of interest. **Nonresponse bias** occurs when respondents and nonrespondents differ in ways relevant to the estimate. A high response rate is desirable but does not guarantee low bias; a lower response rate can still yield useful estimates when nonresponse is understood and adjusted. Professional survey standards therefore examine recruitment, dispositions, weighting and nonresponse patterns, not only a single response-rate percentage.

TARGET POPULATION	SAMPLING FRAME	SELECTED SAMPLE	ACHIEVED SAMPLE

TARGET POPULATION	SAMPLING FRAME	SELECTED SAMPLE	ACHIEVED SAMPLE
Who should the claim cover?	Who can actually be selected?	Who is invited/selected and with what probability or logic?	Who contributes usable evidence after nonresponse, attrition and exclusions?

Figure 3.4. Sampling logic from target population to achieved evidence.
Source: Author's synthesis.

3.12 Measurement, Operationalization, Reliability and Validity

A research design can fail even with perfect sampling if the study does not measure what its conclusions require. **Operationalization** connects an abstract construct to observable indicators or procedures. "Innovation capability," "trust," "export readiness," "corruption exposure" and "customer loyalty" are not directly observable in the same way as invoice value or shipment date. Researchers must explain what evidence will represent each construct and why.

3.12.1 Variables and levels of measurement

A **variable** takes different values across observations. Variables may be categorical or quantitative, observed or latent, continuous or discrete. Measurement levels are commonly described as nominal, ordinal, interval and ratio. Nominal categories have no inherent order; ordinal categories have order without assumed equal distance; interval scales have equal units without a meaningful absolute zero; ratio scales add a meaningful zero. The distinction affects appropriate summaries and models, which Chapter 4 develops.

3.12.2 From construct to indicator

Good operationalization requires content coverage. If "export readiness" is theoretically multidimensional, a single question asking "Are you export ready?" is unlikely to measure it adequately. The researcher can identify dimensions such as product compliance, managerial knowledge, financing, production capacity and market intelligence, then select validated indicators or develop measures through a documented process.

Hinkin (1998) emphasizes systematic scale development for organizational surveys, including item generation, content adequacy, questionnaire administration, item reduction, factor structure and validation. The relevant lesson for undergraduates is that scale development is a research project, not a last-minute wording exercise.

3.12.3 Reliability

Reliability concerns consistency of measurement under defined conditions. **Internal consistency** examines whether items intended to represent a common construct cohere. **Test-retest reliability** examines stability over time where stability is theoretically expected. **Inter-rater reliability** examines agreement among raters when coding or judgment should be reproducible.

Cronbach's alpha (Cronbach, 1951) is widely reported as an internal-consistency coefficient, but it is frequently misinterpreted. A high alpha does not prove unidimensionality or validity, and alpha can increase simply because more similar items are added. Mechanical claims that alpha above a universal threshold automatically validates a scale should be avoided. Reliability must be interpreted with construct structure and purpose.

3.12.4 Validity of measures

Content validity concerns whether indicators adequately cover the construct domain. **Criterion validity** concerns relationships with relevant external criteria, including concurrent or predictive evidence. **Construct validity** concerns whether the pattern of relationships behaves as the theoretical construct should. Convergent evidence asks whether related measures relate appropriately; discriminant evidence asks whether conceptually distinct constructs remain distinguishable.

Measurement validity should be separated from **internal validity** and **external validity** of a study. Internal validity concerns whether an estimated relationship, especially a causal effect, is credible within the study. External validity concerns generalization across populations, settings, treatments or times. A survey scale can be psychometrically valid while a causal inference using it is weak.

3.12.5 Trustworthiness in qualitative research

Qualitative traditions often use the language of **credibility, transferability, dependability and confirmability** rather than translating quantitative validity categories mechanically (Lincoln & Guba, 1985). Credibility can be strengthened through prolonged engagement, triangulation, negative-case analysis, member reflection where appropriate and transparent evidence. Transferability is supported by sufficiently rich contextual description for readers to judge relevance elsewhere. Dependability concerns the traceability of the research process. Confirmability concerns whether interpretations can be connected to evidence rather than hidden researcher preference.

Reflexivity cuts across these criteria. Researchers should examine how their identity, role, access, interests and assumptions shape the inquiry. Reflexivity does not require autobiographical performance. It requires analysis of influences that could materially affect what was asked, disclosed, noticed or interpreted.

3.12.6 Bias as a design problem

Bias is systematic error, not simply disagreement with the researcher. Common threats include:

Table 3.9. *Common sources of bias and design responses.*

Bias / threat	Mechanism	Example	Design response
Selection bias	Included units differ systematically from intended comparison	Only successful exporters agree to interviews about export capability	Broaden recruitment, document refusal, sample contrasting cases
Recall bias	Memory accuracy differs by time, salience or outcome	Failed firms reconstruct early warnings differently years later	Shorter recall periods, records, event anchors, triangulation
Social-desirability bias	Respondents provide normatively acceptable answers	Managers underreport informal payments	Privacy, neutral wording, indirect techniques, alternative data
Common-method bias	Same source/method inflates or structures associations	Employee reports both leadership and own performance in one survey	Multiple sources/times, procedural separation, method diagnostics
Interviewer bias	Interviewer's behavior influences response	Probes only when answers support the hypothesis	Training, guide discipline, reflexive review
Confirmation bias	Researcher preferentially seeks or interprets supportive evidence	Contrary cases are coded as “exceptions” without analysis	Pre-specification, rival explanations, negative-case analysis
Survivorship bias	Only surviving units remain observable	Studying successful internationalizers to infer success factors	Include exits/failures, define cohort prospectively
Attrition bias	Loss to follow-up differs across groups or outcomes	Low-performing firms disappear from panel surveys	Retention plan, attrition comparison, sensitivity analysis
Measurement non-equivalence	Instrument functions differently across groups	“Autonomy” item carries different workplace meaning across cultures	Translation review, cognitive tests, invariance analysis where appropriate

Source/Note: Author’s synthesis from sources cited in the surrounding text.

Podsakoff et al. (2003) provide a widely cited treatment of common method biases in behavioral research. The broader point is that bias control begins in design, not after results are known.

MEASUREMENT QUALITY Do indicators represent the intended construct? Content, criterion and construct evidence; reliability is related but not identical.	INTERNAL INFERENCE Within the study, could bias, confounding, attrition or implementation explain the result?
EXTERNAL INFERENCE To which populations, settings, treatments and time periods can the result reasonably travel?	**QUALITATIVE TRUSTWORTHINESS** Are interpretation, context, reflexivity, negative cases and the analytic trail sufficiently transparent and credible?

Figure 3.5. Measurement and inference quality as distinct but connected tests.
Source: Author’s synthesis.

3.13 Data Management and Analysis Planning Before Data Collection

An analysis plan forces the researcher to explain how observations will become answers. It reduces the risk of collecting interesting but unusable data and helps reveal design weaknesses while they can still be corrected.

3.13.1 Quantitative analysis plans

A quantitative analysis plan should identify the unit of analysis, variables, derived measures, data-quality checks, descriptive summaries, primary comparisons or models, subgroup analyses, handling of missing data, treatment of outliers, sensitivity analyses and criteria for interpreting results. The proposal does not need to reproduce statistical theory from Chapter 4, but it should name methods at the level necessary to establish fit.

A strong plan links each method to a question. If the question is descriptive, a causal model may be unnecessary. If the question concerns change over time, a single cross-sectional mean comparison is insufficient. If the treatment is randomized by branch, analysis should recognize branch-level assignment. If the outcome is a count or category, the planned model should respect the outcome structure.

3.13.2 Qualitative analysis plans

Qualitative analysis should be described beyond "responses will be analyzed thematically." The researcher should state the analytic tradition, coding orientation, whether coding is inductive, deductive or hybrid, how codes will be developed and revised, how themes or categories will be constructed, how divergent cases will be handled, what role memos or matrices will play, and how interpretation will be checked against the dataset.

Braun and Clarke's thematic-analysis approach treats themes as analytic constructions rather than entities that "emerge" automatically. Gioia-style analysis emphasizes a transparent data structure in inductive organizational research. Grounded theory uses constant comparison and theoretical sampling. Content analysis may use structured categories and sometimes frequencies. The plan should reflect the actual method rather than generic qualitative vocabulary.

3.13.3 Software is a tool, not an analytical method

Spreadsheet software, R, Python, Stata, SPSS and SAS can support quantitative analysis. NVivo, ATLAS.ti and MAXQDA can support qualitative coding and retrieval. Reference managers such as Zotero, EndNote and Mendeley can support source organization. Git or other version-control systems can improve reproducibility for code. None of these tools determines the validity of the study.

A researcher should select software based on required functionality, reproducibility, team competence, licensing, data sensitivity and institutional support. Excel may be entirely adequate for a small descriptive dataset but fragile for complex, repeatable transformations. R or Python can support scripted reproducibility but require disciplined coding. Proprietary software can be highly capable while creating access barriers for replication.

3.13.4 Data-management plans

A data-management plan describes file naming, variable definitions, identifiers, access controls, backup, encryption where appropriate, linkage keys, retention, destruction, version control and responsibilities. A codebook should define each variable, source, units, permissible values, missing-value codes and transformations. Qualitative projects should document transcription conventions, pseudonymization and the relationship between consent materials and stored data.

The FAIR principles encourage research outputs to be findable, accessible, interoperable and reusable when appropriate (Wilkinson et al., 2016). "Accessible" does not mean all data should be public. Privacy, commercial confidentiality, legal restrictions, indigenous or community interests and contractual duties can justify controlled or closed access. Responsible open science balances transparency with legitimate protection.

Data minimization and provenance should be planned together. Collecting fewer identifiers reduces privacy exposure, while preserving source metadata, transformation logs and codebooks makes the evidential chain auditable. A well-governed project should be able to distinguish raw source material from cleaned or derived files, identify who changed what and why, and reproduce the pathway from original evidence to analytic dataset. When a correction is necessary, versioned records should preserve the history rather than silently overwrite it.

Table 3.10. *Data-management planning across quantitative and qualitative research.*

Planning element	Quantitative example	Qualitative example	Integrity purpose
File structure	Raw, cleaned and analysis datasets separated	Audio, transcripts, de-identified transcripts, coding project	Preserve provenance
Codebook / data dictionary	Variable name, label, unit, missing code, source	Participant ID, interview date, context, transcript status	Reduce ambiguity
Version control	Scripted transformations with dated releases	Memo and codebook revision history	Make changes traceable
Access control	Restricted folder for identifiable linkage key	Separate identity file from transcripts	Protect confidentiality
Analysis record	Reproducible scripts and output log	Audit trail of coding and theme revisions	Support transparency
Retention / destruction	Defined by law, consent and institutional policy	Same, including audio deletion if promised	Keep commitments to participants

Source/Note: Author's synthesis from sources cited in the surrounding text.

EVIDENCE CHECK | Reproducibility begins before analysis
A researcher cannot reconstruct provenance after raw data have been overwritten, variable meanings forgotten or interview files renamed inconsistently. Data management is part of methodology because it affects whether evidence remains interpretable, auditable and ethically protected.

3.14 Research Ethics, Data Protection and Research Integrity

Research ethics is not reducible to obtaining approval from an ethics committee. Approval is an institutional judgment that a proposed study meets applicable requirements at a point in time. Researchers remain responsible for what they actually do in the field, how they protect participants, how they respond to new risks and how honestly they report results.

3.14.1 Foundational ethical principles, respect for persons and informed consent

Research ethics draws on several overlapping traditions. The Belmont Report, developed for human-subjects research, articulates respect for persons, beneficence and justice as foundational principles and connects them to informed consent, risk-benefit assessment and fair participant selection (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). Its U.S. institutional origin does not make it a universal legal code, but the framework remains useful for identifying questions that business and social researchers must answer alongside applicable national law, institutional policy and disciplinary standards.

Informed consent requires more than a signed form. Participants should receive information that a reasonable person would need to decide whether to participate: the purpose and procedures, expected time, foreseeable risks and benefits, voluntary nature of participation, confidentiality limits, data use, relevant compensation, contact information and the right to decline or withdraw subject to legitimate limits already disclosed.

Consent is especially sensitive in hierarchical settings. Employees may fear that refusal will be reported to management. Suppliers may worry about future contracts. Students may think participation affects assessment. Small traders dealing with officials may fear regulatory consequences. Researchers should design recruitment so that authority figures do not unnecessarily know who accepted or declined, and should avoid incentives so large that they undermine voluntariness in context.

Organizational permission is not participant consent. A chief executive can authorize access to a workplace but cannot ordinarily consent on behalf of employees to disclose their private experiences. Similarly, public availability of a social-media post does not automatically eliminate all ethical duties regarding aggregation, re-identification or contextual harm.

3.14.2 Risk, benefit and proportionality

Research risks can be physical, psychological, social, economic, professional, legal or reputational. In business research, the most serious risks are often not physical. An interview about corruption can expose a participant to retaliation. A dataset linking credit behavior to identity can create financial harm. A publication that identifies a small firm through contextual details can damage relationships even after names are removed.

Researchers should minimize foreseeable risk, collect no more identifying data than necessary, and ensure that risk is proportionate to the value and design of the inquiry. “Minimal risk” should not be assumed merely because the study uses an online survey.

3.14.3 Confidentiality, anonymity and pseudonymity

Anonymity means identities are not known or reasonably linkable to the research data under the study design. **Confidentiality** means identities may be known to the research team but are protected from unauthorized disclosure. **Pseudonymization** replaces direct identifiers with codes while preserving a separate linkage possibility. Researchers should not promise anonymity when email addresses, consent forms, IP addresses or longitudinal identifiers permit re-identification.

Confidentiality also has limits. Legal duties, safeguarding concerns or court orders may override promises in some jurisdictions and settings. Consent language should describe material limits rather than offer impossible guarantees.

3.14.4 Data protection in Rwanda and cross-border research

For research involving personal data in Rwanda, Law No. 058/2021 of 13/10/2021 relating to the protection of personal data and privacy, officially gazetted on 15 October 2021, is a central legal reference. The law designates the National Cyber Security Authority (NCSA) as the supervisory authority. Current implementation and compliance services are provided through the Data Protection and Privacy Office (DPO). Among other requirements, the framework addresses

lawful and transparent processing, security, accountability, data-subject rights, controller and processor obligations, and cross-border storage and transfer (Data Protection and Privacy Office [DPO], n.d.; Republic of Rwanda, 2021).

A research protocol should therefore map the full personal-data lifecycle before collection begins: what data are necessary, the lawful basis and research purpose, who determines the means and purposes of processing, which processors or cloud vendors are involved, where the data will be stored, who can access them, how long they will be retained, and how data-subject rights can be honored. As of August 2026, the DPO maintains specific authorization processes for storing and transferring personal data outside Rwanda. Article 48 of the law also identifies specified circumstances in which cross-border transfer may occur, while Article 50 requires authorization for storage outside Rwanda. Researchers using cross-border cloud infrastructure should therefore verify the current legal basis, authorization, contractual safeguards and security requirements rather than assuming that institutional ethics approval alone authorizes international processing (DPO, n.d.; Republic of Rwanda, 2021).

The African Union Data Policy Framework (African Union, 2022) seeks to strengthen data governance and trusted data use across the continent, including digital trade and research. Regional interoperability can reduce friction and support collaboration, but harmonization is not self-validating. Cross-border data arrangements should also be evaluated for national law, sovereignty, institutional capacity, accountability, power asymmetry, data localization requirements, cybersecurity and whether participants were given an intelligible account of where their information may travel. Legal compliance is a minimum condition; responsible stewardship also asks whether the proposed data architecture is necessary and proportionate to the research purpose.

POLICY LENS | "Cloud-based" is a data-governance decision
Uploading interview transcripts or identifiable survey data to a cloud service can create a cross-border transfer, vendor-access and retention problem. Before using any platform, researchers should know where data are processed, whether contractual terms permit secondary use, what security controls exist, and whether the transfer complies with applicable law and consent. Convenience does not suspend legal or moral responsibility.

3.14.5 Vulnerable participants and contextual vulnerability

Vulnerability is not a permanent label attached to categories of people. It can arise from dependence, limited alternatives, legal status, economic insecurity, literacy, age, disability, trauma, power relations or the sensitivity of the topic. A small supplier negotiating with a dominant buyer may be vulnerable in one study because disclosure could affect livelihood. Researchers should identify the mechanism of vulnerability and design protections accordingly.

3.14.6 Deception and incomplete disclosure

Some experimental questions may require withholding limited information to avoid demand effects. Deception raises a high ethical burden. Researchers should ask whether the question can be answered without it, whether risks are minimal, whether deception concerns material participant interests, and whether debriefing is necessary and effective. Business settings involving employees or customers require additional caution because organizations can possess unequal power and extensive behavioral data.

3.14.7 Research integrity: fabrication, falsification and plagiarism

Fabrication invents data or sources. **Falsification** manipulates research materials, processes or data so the record no longer represents what occurred. **Plagiarism** presents another's words or ideas without appropriate attribution. These are clear violations, but integrity failures also occur in subtler forms: selective exclusion of inconvenient observations, undisclosed outcome switching, misleading image or chart manipulation, HARKing (presenting post-hoc hypotheses as pre-specified), inappropriate authorship, duplicate publication, citation padding and suppression of sponsor-unfriendly findings.

Contemporary integrity frameworks broaden the responsibility beyond avoiding fabrication, falsification and plagiarism. The revised European Code of Conduct for Research Integrity organizes good practice around reliability, honesty, respect and accountability and applies those principles to research environments, data practices, collaboration, publication and review (All European Academies [ALLEA], 2023). These principles are useful internationally as professional guidance, but they do not displace national law, disciplinary standards or institutional procedures. Their value lies in making integrity a property of the whole research process rather than a narrow misconduct threshold.

Integrity is therefore about the entire evidential record. A result can be technically "real" yet misleading if the researcher reports only the analysis that produced significance after trying many alternatives. Pre-specification, registered reports, reproducible scripts, transparent sensitivity analyses and clear distinctions between confirmatory and exploratory work can reduce such risks (Munafò et al., 2017; Nosek et al., 2018).

3.14.8 Conflicts of interest and sponsor influence

Business research often depends on organizational access or funding. Sponsorship does not invalidate a study, but material interests should be disclosed and governance should protect analytic independence. Contracts should clarify who owns data, whether the researcher can publish adverse findings, who reviews outputs for confidentiality versus editorial control, and how disagreements are handled.

A client may legitimately correct factual errors or identify confidential information. A client should not be allowed to rewrite findings merely because the evidence is commercially uncomfortable. If publication rights are constrained, the researcher should disclose those constraints to the relevant audience.

3.14.9 Ethics review and continuing responsibility

Ethics committees or institutional review boards review study risks, consent and protections, but their jurisdiction varies. Some secondary-data work may be exempt or not involve human participants, yet privacy, contractual and integrity responsibilities remain. Researchers should not infer "no ethics application required" means "no ethical issues exist." Material changes to recruitment, instruments, data sources or risk normally require consultation or amendment under applicable institutional rules.

3.14.10 Authorship, contributorship and equitable collaboration

Authorship is both credit and accountability. Researchers should agree contribution expectations early, revisit them as work changes, and ensure that every listed author has made a defensible scholarly contribution and accepts responsibility appropriate to that contribution. Gift authorship, ghost authorship and coercive authorship distort the research record because they misallocate both credit and responsibility. Where journals permit contributorship statements, explicit role descriptions can improve transparency, but a taxonomy cannot substitute for substantive judgment about who actually contributed to the intellectual work.

Cross-border collaboration adds a further question of fairness. The Cape Town Statement links research integrity with fairness and equity, emphasizing shared agenda setting, recognition of local knowledge, transparent budgeting, equitable data access and appropriate credit in collaborations involving resource asymmetries (Horn et al., 2023). In African business and trade research, local collaborators should not be reduced to recruiters, translators or data collectors when they contribute contextual knowledge, design judgment or interpretation. Conversely, local affiliation alone does not create automatic authorship. Credit should follow genuine contribution, with power imbalances actively prevented from determining the byline.

Collaboration agreements should therefore address decision rights over protocols, data custody and reuse, publication review, confidentiality, intellectual property, authorship, dispute resolution, capacity development and the return of useful findings to participating organizations or communities. These agreements protect both scholarly independence and legitimate commercial or community interests. They are especially important when the institution that funds the research is not the institution that bears the fieldwork burden or supplies the data.

Table 3.11. *Research ethics: minimum technical questions and deeper moral questions.*

Ethical issue	Minimum technical question	Deeper moral question
Consent	Was valid consent obtained under applicable rules?	Could the person genuinely refuse without fear or undue pressure?
Confidentiality	Are identifiers protected?	Could contextual details still expose the participant to foreseeable harm?
Data minimization	Is the field legally permissible to collect?	Do we actually need it for the question, or are we accumulating power without purpose?
Incentives	Is compensation allowed?	Does the amount distort voluntariness in this economic context?
Sponsor review	Is the agreement contractually clear?	Can inconvenient truth be reported without retaliation?
AI assistance	Is tool use permitted?	Has human responsibility been displaced or confidential material exposed?
Publication	Are reporting requirements met?	Have uncertainty, adverse findings and affected voices been represented fairly?
Authorship / collaboration	Are contributions, rights and responsibilities documented?	Are credit, data access, decision power and local knowledge being handled fairly rather than by hierarchy?

Source/Note: Author's synthesis from sources cited in the surrounding text.

3.15 Responsible Use of Generative AI in Research

Generative AI can assist research workflows, but fluency must never be mistaken for evidence. Large language models can summarize, classify, brainstorm, translate, draft code and help re-express text. They can also fabricate references, misstate methods, reproduce biases, expose confidential information, and generate plausible analysis that the researcher does not understand.

UNESCO's guidance on generative AI in education and research emphasizes a human-centred approach, data privacy and institutional capacity (Miao & Holmes, 2023). The key research principle is straightforward: **the accountable researcher remains responsible for every claim, source, method, analysis and disclosure, regardless of which tool assisted the workflow.**

Because model capabilities, provider terms, data-retention settings and institutional policies change rapidly, responsible practice cannot rely on a static list of approved tools. Where AI assistance materially affects a research workflow, the researcher should record the tool and version when identifiable, the date and purpose of use, the type of data sent to the system, the verification performed, and any disclosure required by the institution, funder or publisher. The objective is auditability and accountable human judgment, not performative disclosure of trivial spell-checking or formatting assistance.

That accountability has become more consequential as AI use has spread. Resnik and Hosseini (2026) argue that hallucinated citations can, in some circumstances, rise beyond poor scholarship and potentially constitute research misconduct when fabricated citations function as data supporting research findings. The precise legal classification varies by jurisdiction, but the methodological duty is not ambiguous: every citation must exist, must be read or otherwise verified by the accountable researcher, and must actually support the proposition for which it is cited. Fluency, citation-like formatting and confident language are never substitutes for source verification.

3.15.1 Legitimate research assistance

Potentially legitimate uses include generating alternative search terms, improving code comments, checking grammar, converting documented code between languages, suggesting interview probes for researcher review, organizing non-sensitive notes, or helping draft a data dictionary. These uses still require verification. AI-generated search terms may systematically omit relevant traditions; code may run while implementing the wrong estimand; summaries may invent emphasis not present in the source.

3.15.2 High-risk and unacceptable substitutions

A researcher should not use AI to fabricate participants, fabricate quotations, invent citations, manufacture observations, simulate “evidence” and present it as collected data, or write a literature review from references that the researcher has not verified. Synthetic data can have legitimate methodological uses, but it must be labelled as synthetic and should not be represented as evidence about real people or firms.

Uploading confidential transcripts, proprietary client data, unpublished manuscripts or personally identifiable information into an external model can breach consent, contracts or data-protection law. “The tool does not remember” is not a sufficient governance statement unless the actual service terms and technical controls support it.

3.15.3 AI-assisted qualitative analysis

AI can help retrieve text segments, suggest candidate codes or compare coding consistency, but interpretation remains theory- and context-dependent. The researcher should know what material was sent to the tool, whether data were de-identified, what model/version was used if materially relevant, and how human review corrected errors. An AI-produced theme is not a finding until it is examined against the corpus and justified analytically.

3.15.4 AI-assisted quantitative analysis

Code generation can accelerate statistical work, but generated code may silently change reference categories, drop missing observations, apply the wrong estimator or produce charts with misleading defaults. Reproducible scripts should be inspected line by line, tested on known cases and linked to documented data transformations. Researchers who cannot explain an analysis should not submit it merely because software executed successfully.

> **PRACTICE TOOL | The V-PAD control for research AI**
> **V - Verify:** check every factual claim, citation, calculation and code path against authoritative sources or reproducible tests. **P - Protect:** never expose personal, confidential or restricted data without lawful and ethically justified controls. **A - Attribute:** disclose material AI assistance when required by institutional, publisher or project rules and never attribute AI-generated text as evidence. **D - Defend:** the human researcher must be able to defend every methodological decision and conclusion without saying “the model produced it.”

3.16 Research Proposals and Project Management

A research proposal is a design argument. It should persuade a critical reader that the problem matters, the question is answerable, the design fits the question, the evidence can realistically be obtained, the analysis is appropriate, the risks are managed and the project can be completed with available time and resources.

3.16.1 Core proposal architecture

A strong undergraduate business or trade proposal usually includes the title; background and problem statement; purpose and objectives; research questions or hypotheses; significance and scope; literature synthesis; theoretical or conceptual framework; methodology and design; setting, population and sampling; instruments or data sources; data-collection procedures; analysis plan; quality or validity strategy; ethics and data management; limitations; workplan; budget where relevant; references; and appendices containing draft instruments or permissions.

Institutional templates vary. The intellectual logic should not. A proposal with beautifully formatted sections but inconsistent questions, variables and design is not ready for approval.

3.16.2 Feasibility

Feasibility is part of rigour. A design requiring confidential bank-level microdata is not strong if access is speculative. A four-country ethnography is not defensible within six weeks. A survey translated into four languages requires time and resources for instrument equivalence. A quasi-experimental study needs a credible comparison and sufficient pre/post data before it is promised.

Feasibility assessment should include access, recruitment, sample availability, travel, permissions, technology, fieldworker capability, security, transcription, data quality, analysis skills and contingency plans. The best design is the strongest design that can be executed ethically and competently, not the most ambitious design that can be described.

3.16.3 Work breakdown and Gantt logic

A project plan should identify dependencies. Ethics approval must precede recruitment. Translation precedes piloting. Data collection should not begin before the instrument and sampling frame are finalized. Cleaning and transcription can sometimes proceed in parallel with fieldwork, but this should be planned.

Table 3.12. *Research project phases, dependencies and quality gates.*

Phase	Core outputs	Dependency	Typical quality gate
Problem and evidence review	Approved question, preliminary framework	Access to literature and stakeholder problem definition	Research gap and scope are defensible
Design and proposal	Methodology, sampling, instruments, analysis plan	Problem alignment	Supervisor/method review completed
Ethics and permissions	Approval, data agreements, organizational access	Finalized procedures	No recruitment before required approvals
Pilot	Tested instrument/protocol	Ethics/permission as applicable	Major ambiguities corrected and documented
Main evidence collection	Complete field or secondary dataset	Frame/instrument readiness	Field monitoring and protocol adherence
Data preparation	Clean dataset/transcripts/codebook	Evidence collection	Raw data preserved, exclusions logged
Analysis	Results linked to research questions	Quality checks	Pre-specified analyses plus documented exploration
Interpretation and writing	Findings, limitations, implications	Analysis complete enough	Claims do not exceed evidence
Dissemination and archiving	Final report, presentation, repository package as appropriate	Approval/confidentiality review	Privacy, attribution and correction procedures satisfied

Source/Note: Author's synthesis from sources cited in the surrounding text.

3.16.4 Risk register

Research plans should anticipate risks that can compromise validity or ethics: frame failure, low response, data-loss, participant distress, security disruption, sponsor interference, staff turnover, political sensitivity, translation failure or software access. Each risk should have probability, impact, mitigation, trigger and owner. Risk management does not eliminate uncertainty; it makes response responsibilities explicit.

The register should distinguish risks to project delivery from risks to inference and participant welfare. A delayed interview schedule is a project risk; systematic loss of small firms from the sampling frame is an inferential risk; interviewing employees about misconduct in a room visible to managers is an ethical and safety risk. Some risks belong

in more than one category. The register should be reviewed at predefined milestones and whenever field conditions materially change, with residual risk and any protocol amendment documented rather than treated as informal project memory.

3.17 Dissemination, Transparency, Reproducibility and Responsible Openness

Research is incomplete when evidence cannot be understood by its intended audience. Dissemination should match the claim and stakeholder. An academic article may foreground theory and method; a policy brief may foreground decision implications; a company report may foreground operational recommendations; a public communication may require accessible explanation and privacy protection. Chapter 2 provides the communication skills; Chapter 3 adds methodological obligations.

3.17.1 Reporting guidelines

Reporting guidelines can improve transparency by reminding researchers to disclose information readers need. PRISMA 2020 addresses systematic reviews. COREQ provides a checklist for reporting interviews and focus groups in health research and is often consulted more broadly for qualitative transparency (Tong et al., 2007). STROBE addresses observational epidemiological research (von Elm et al., 2007). Discipline-specific guidelines should be used intelligently rather than copied into unrelated designs. The principle is transparency about design, not bureaucratic checklist completion.

3.17.2 Preregistration and registered reports

Preregistration records hypotheses, outcomes and analysis plans before observing the relevant results. It can clarify which analyses were confirmatory and which were exploratory, reducing undisclosed flexibility. Preregistration does not make a poor theory good, prevent all bias or forbid learning from data. Deviations can be legitimate if they are disclosed.

Registered reports move peer review earlier by evaluating the research question and method before results are known, with publication decisions less dependent on whether findings are statistically significant. Such models are not available in every business journal but illustrate a broader principle: strong inference benefits when design quality is evaluated before outcomes influence judgment.

3.17.3 Replication materials

Where legal and ethical constraints allow, researchers can share questionnaires, interview guides, codebooks, analysis scripts, synthetic examples, metadata and de-identified data. Replication does not always require public raw data. Restricted-access repositories, safe rooms or controlled agreements may be necessary for sensitive business data.

Researchers should also distinguish computational reproducibility from broader replicability. Computational reproducibility asks whether the reported result can be regenerated from the same data and analytic procedures. Replicability asks whether a new study, using new observations and a sufficiently comparable design, reaches a conclusion compatible with the original claim. Robustness asks whether reasonable alternative specifications, coding decisions or assumptions materially change the result. Terminology varies across disciplines, so authors should define what they mean and provide the materials necessary for the specific form of verification they claim to support.

A data-availability statement should therefore say what can be shared, where it can be accessed, what restrictions apply and why. “Data available on request” is weak when no access conditions, responsible contact or legal basis are specified. Conversely, refusing public release can be methodologically and ethically correct when disclosure would violate consent, commercial confidentiality, data-protection law or legitimate community interests. Responsible openness is transparent about both access and justified limits.

3.17.4 Corrections and negative results

Integrity includes correcting errors after publication and reporting null or adverse findings when they are relevant. A study does not fail because an expected effect is absent. It fails methodologically when design weaknesses make the absence uninterpretable or when the researcher suppresses inconvenient evidence. Organizations also need correction cultures in which decision-relevant findings can be revised without treating revision as personal defeat.

Null findings should be interpreted against design sensitivity, confidence intervals, measurement quality and the effect size that would matter for the decision rather than translated mechanically into “no effect.” Adverse or unexpected findings may be the most decision-relevant result of an evaluation and should not be buried because they complicate a sponsor narrative. When material errors are discovered after dissemination, the responsible response is proportionate correction of the record, ranging from an amended report or corrigendum to retraction where the central findings are no longer trustworthy.

3.18 Designing Research for African Business and Trade Contexts

African research should neither be treated as an exotic exception nor forced into methods developed elsewhere without contextual examination. The continent contains diverse legal systems, languages, infrastructure, market structures, levels of formality, data institutions and histories. Methodological quality therefore requires attention to the actual setting rather than continent-wide generalization.

3.18.1 Informality and frame construction

In sectors with substantial informality, business registers may omit many economically active units. Researchers can use area-based sampling, venue-based approaches, adaptive cluster methods, trader associations, market enumeration or mixed frames when appropriate. Each approach changes the universe of inference. A study of registered exporters cannot be generalized to informal cross-border traders without evidence that the populations are comparable.

3.18.2 Multilingual research

Translation may involve national languages, regional lingua francas and technical business terminology. Researchers should record which language was used in each interview, who translated, whether transcripts preserve original expressions, and how meaning was adjudicated. Translation is part of analysis when a concept has no exact equivalent.

3.18.3 Infrastructure and digital exclusion

Online surveys can appear cheap and scalable but systematically exclude participants with weak connectivity, limited data budgets, device constraints or low digital literacy. Mixed-mode designs can reduce exclusion but introduce mode effects. The researcher should distinguish the convenience of digital collection from representativeness.

3.18.4 Institutional trust and respondent risk

Questions about tax, corruption, informality, political connections or border payments can create genuine fear. A participant may rationally withhold information if confidentiality protections are weak or researchers are perceived as connected to authorities. Trust cannot be manufactured by adding an academic logo to a consent form. Local partnerships, clear independence, data minimization and realistic promises matter.

3.18.5 Sovereignty and extractive research

Cross-border research partnerships can reproduce unequal control when external institutions define questions, control budgets and data, receive authorship, and leave local researchers or communities with little durable benefit. Methodological excellence therefore includes governance of knowledge production. The Cape Town Statement explicitly links fairness and equity to research integrity and warns against agenda setting, authorship, data-access and resource practices that systematically disadvantage lower-resourced partners (Horn et al., 2023). Questions about data ownership, analytical capacity, authorship, benefit sharing, intellectual property and long-term access should therefore be negotiated at the design stage rather than assumed after data collection.

The African Union's data-governance agenda offers opportunities for interoperable, rights-respecting research infrastructure. Yet continental coordination should be implemented with attention to subsidiarity, national law, institutional capacity and the risk that centralization can move data power farther from the people who generated the information. Research infrastructure is never morally neutral simply because it is technically efficient.

3.18.6 Researcher safety, reciprocal value and field logistics

Business and trade fieldwork can expose researchers and field staff to risks that are easily overlooked in classroom protocols. Studies of informal payments, smuggling, labor abuse, political connections, sanctions, conflict-affected trade or commercially sensitive supply chains may create surveillance, retaliation, legal or physical risks for both participants and researchers. A field plan should therefore specify safe interview locations, communication and escalation procedures, transport and lone-working arrangements where relevant, handling of devices and notes, emergency contacts, and clear stopping rules when conditions become unsafe. No data point is valuable enough to justify foreseeable serious harm to a participant or fieldworker.

Reciprocity does not require promising direct benefits that the study cannot deliver, nor should compensation purchase preferred answers. It does require respect for participant time and local expertise and, where feasible, a plan to return findings in a form that participating organizations or communities can use. In cross-border collaborations, reciprocal value can also include shared analytical capacity, access to cleaned datasets under lawful conditions, co-development of instruments and genuine participation in interpretation. These practices strengthen both justice and methodological validity because they reduce extractive relationships that can distort trust, disclosure and meaning.

AFRICAN CONTEXT | Local knowledge is evidence, not decoration
Contextual expertise should enter problem formulation, sampling, interpretation and ethics, not only the case-study section of a report. A technically sophisticated design that misunderstands how firms register, how traders route transactions, how languages are used or how participants perceive state authority can generate precise answers to the wrong question.

3.19 African Case Study: Evaluating Nigeria's YouWiN! Entrepreneurship Programme

McKenzie (2017) provides an unusually instructive African business-research case because the study combines a consequential policy question, experimental design, long follow-up and outcomes directly relevant to enterprise growth. The national competition made large grants available to business-plan applicants. Random assignment among a relevant group of applicants enabled the evaluation to distinguish selection into winning from the causal effect of winning.

3.19.1 The research problem

Nigeria, like many economies, has many very small firms and relatively few firms that grow into substantial employers. Policymakers therefore face two linked uncertainties: whether high-growth entrepreneurial potential can be identified, and whether relieving capital constraints can convert that potential into business growth. Simply comparing selected winners with rejected applicants would not answer the second question because judges intentionally selected stronger plans.

3.19.2 The design solution

Randomization created a comparison among applicants whose treatment status was not determined solely by merit. The evaluation could therefore estimate the effect of receiving the programme grant among the population subject to random assignment. Follow-up surveys tracked outcomes over five years, allowing the researcher to observe entry, survival and growth rather than only immediate spending.

3.19.3 Measurement choices

Employment, profits, sales and firm survival are economically meaningful outcomes, but each has measurement challenges. Employment can be seasonal or informal. Profit reporting can be noisy in small firms with limited accounts. Survival depends on how “operating” is defined. Longitudinal follow-up creates attrition risk. A rigorous evaluation therefore needs consistent definitions, respondent tracking, validation where feasible and transparent treatment of missing outcomes.

3.19.4 What the evidence supports

The study reports that winning the competition increased firm entry, survival, profits, sales and employment and raised the probability of operating a firm with ten or more workers by more than 20 percentage points (McKenzie, 2017). The randomization strengthens the causal interpretation for the eligible applicants and programme context.

The study does **not** prove that grants are always the optimal entrepreneurship policy. External validity depends on applicant quality, grant size, screening, complementary support, macroeconomic conditions and programme implementation. Scaling can change administrative quality and participant composition. Opportunity cost also matters: the relevant policy comparison may be grants versus credit guarantees, infrastructure, tax reform, managerial training or no intervention.

3.19.5 Ethical and policy questions

Scarce grant allocation creates distributive questions even when randomization is methodologically useful. Was random assignment fair among applicants who had met eligibility thresholds? Did participants understand the research component? Were unsuccessful applicants exposed to burdens without realistic benefit? Could political or administrative interests influence treatment integrity? These are not arguments against experimentation. They are reminders that causal identification and moral legitimacy are separate tests.

Table 3.13. *Methodological audit of the Nigeria YouWiN! evaluation.*

Design element	YouWiN! lesson	Transferable methodological principle
Practical problem	Few firms scale into larger employers	Define the decision problem before selecting metrics
Selection problem	Strong applicants are more likely to win	Selection into treatment can mimic treatment effect

Design element	YouWiN! lesson	Transferable methodological principle
Identification	Random assignment among eligible applicants	Design the comparison before analyzing outcomes
Time	Five-year applicant tracking	Important outcomes may require longitudinal follow-up
Outcome quality	Entry, survival, sales, profit, employment	Use multiple outcomes but define and measure each consistently
External validity	Nigerian competition applicants in a specific period	State the population and institutional conditions of inference
Policy interpretation	Positive effects do not settle opportunity cost	Causal effect is one input into policy judgment, not the whole decision

Source/Note: Author's synthesis from sources cited in the surrounding text.

MANAGERIAL INSIGHT | The case changes the question from "Did winners succeed?" to "What did winning change?"
This distinction is the heart of causal research. Whenever managers compare trained with untrained staff, digital adopters with non-adopters, exporters with non-exporters or funded with unfunded firms, they should ask whether pre-existing differences could explain the observed outcome. Research design exists to make such alternative explanations visible and, where possible, weaker.

3.20 Comparative Global Case: Management Practices in Indian Textile Firms

Bloom, Eifert, Mahajan, McKenzie, and Roberts (2013) asked whether management practices causally affect productivity in large Indian textile firms. The problem was methodologically difficult because better-managed firms could also differ in capital, owners, workforce quality or market opportunities. Cross-sectional comparisons between management scores and productivity would therefore be vulnerable to confounding.

The researchers implemented a field experiment in which randomly chosen treatment plants received intensive consulting on management practices and were compared with control plants. The study reported that adoption of the recommended management practices increased productivity by about 17% in the first year through improvements including quality, efficiency and reduced inventory, and that treated firms expanded more over time (Bloom et al., 2013).

The study illustrates several design principles. First, the intervention had to be specified with sufficient fidelity to know what "management improvement" meant. Second, treatment assignment and implementation occurred at organizational levels where spillovers could be possible. Third, operational outcomes were preferable to self-reported satisfaction alone. Fourth, a real-firm field experiment gains realism but sacrifices the scale of a national survey. External validity therefore depends on how similar other firms are to the studied textile organizations.

Comparing the Indian and Nigerian studies also reveals that "experiment" is not a single method. YouWiN! randomized access to capital within an entrepreneurship policy; the Indian study randomized management consulting among established textile plants. The units, treatments, outcomes, implementation risks and generalization boundaries differ. Random assignment strengthens internal causal inference, but it does not remove the need to understand context, mechanism or scalability.

Table 3.14. *Comparing the Nigeria and India field experiments.*

Dimension	Nigeria YouWiN!	India management experiment	Research-design lesson
Unit	Entrepreneur/applicant and resulting firm	Textile plant/firm	Unit must match the causal claim
Intervention	Business grant through competition	Intensive management consulting	Treatment must be operationally defined
Central threat addressed	Selection of stronger entrepreneurs	Better-managed firms differ systematically from others	Randomization addresses baseline confounding in expectation
Outcome horizon	Multi-year entry, survival, growth	Productivity and later expansion	Time horizon must fit the mechanism
External-validity boundary	Competition applicants and policy setting	Large Indian textile firms	Internal validity and generalizability are distinct
Managerial implication	Capital can unlock growth for selected entrepreneurs	Management practices can causally affect performance	Evidence should inform, not mechanically dictate, replication

Source/Note: Author's synthesis from sources cited in the surrounding text.

3.21 Faith, Ethics and Moral Reasoning: Research as Truthful Stewardship

Research is a technical practice, but it is also a moral practice because it concerns truth, power, persons and consequences. Researchers decide whose questions matter, whose experiences count as evidence, which uncertainty is disclosed, what information is collected, what risks participants bear and how results are represented. Christian moral reasoning therefore enters research methods at the level of purpose and conduct, not as a devotional sentence added after methodology.

3.21.1 Truth is not a negotiable research output

Christian Scripture treats truthfulness as a moral obligation. Proverbs 12:22 condemns lying lips, while Proverbs 11:1 condemns dishonest scales. The latter is especially relevant to measurement: a technically impressive instrument used to distort rather than clarify is a modern dishonest measure. Research fabrication, selective reporting and misleading certainty violate the same moral principle because they present an unfaithful representation of reality.

Proverbs 18:17 observes that the first account can appear right until it is examined. This captures a core epistemic discipline: claims deserve testing, rival explanations and corroboration. 1 Thessalonians 5:21 calls believers to test all things and hold on to what is good. This is not a biblical endorsement of any specific statistical method. It is a moral orientation toward discernment rather than credulity.

3.21.2 Human beings are not merely data points

Christian anthropology grounds human dignity in the created status of persons rather than their market value, productivity or usefulness (Genesis 1:26-27). Research participants therefore cannot be treated as raw material for publications. Consent, confidentiality and minimization of harm are not bureaucratic obstacles to knowledge production; they express respect for persons whose lives exceed the variables recorded about them.

Neighbour-love (Mark 12:31) requires researchers to ask how they would want their own information, vulnerabilities and words handled under similar power conditions. This principle does not imply that participants control the conclusions or that researchers must avoid unwelcome findings. Love of neighbour is compatible with truthful correction. It prohibits exploitation and deception for professional gain.

3.21.3 Stewardship of evidence and entrusted access

Researchers are stewards of entrusted information. 1 Corinthians 4:2 identifies faithfulness as a requirement of stewardship. Organizational access, confidential records, public funding and participant time are entrusted resources. Researchers should collect only what they can protect and use responsibly, preserve records accurately, and refuse pressure to manipulate results for sponsors or ideological conformity.

Stewardship also applies to public resources. A methodologically weak study can waste money, participant time and policy attention. Rigour is therefore not merely academic prestige. It is responsible use of resources.

3.21.4 Justice in sampling, interpretation and burden

Justice asks who bears research burdens and who receives benefit or visibility. Convenient populations can be over-researched because they are accessible, while marginalized groups remain absent from evidence or are included only as objects of deficit narratives. African researchers and participants can provide data while external institutions receive authorship and analytical authority. Justice requires fair inclusion, but not quota symbolism. The study must include groups because the question and consequences make their evidence relevant.

Justice also governs interpretation. Researchers should not generalize a finding from one city to “Africa,” treat informality as moral deficiency, or describe communities through stereotypes when data show institutional adaptation. Honest contextualization is a form of respect.

3.21.5 Alternative ethical perspectives

Christian ethics should engage other serious traditions fairly. **Deontological and rights-based ethics** emphasize duties, consent and protection against using persons merely as means. **Consequentialist ethics** asks whether expected benefits justify risks and whether harms can be reduced. **Virtue ethics** emphasizes the character of the researcher, including honesty, courage, humility and practical wisdom. **Care ethics** emphasizes relationships, dependency and the moral significance of responsiveness to vulnerability. Professional research-integrity frameworks emphasize transparency, accountability and reliability.

These traditions overlap with Christian moral concerns in important ways while differing in foundations and conflict-resolution principles. A Christian researcher should not claim that consent or data minimization are uniquely Christian inventions. The distinctive theological claim is that truth and human dignity are accountable to God and do not depend solely on institutional consensus, social utility or professional reputation.

3.21.6 Moral courage under sponsor pressure

A researcher may discover that a favoured intervention does not work, a client has misrepresented outcomes, or a supervisor prefers a more "positive" conclusion than the evidence supports. Technical skill is insufficient at this point. Moral courage may require refusing to fabricate, documenting disagreement, escalating concerns or withdrawing from a project whose terms make truthful reporting impossible.

Acts 5:29 articulates a theological principle of ultimate accountability to God above human authority. Applied carefully, it reminds Christian professionals that institutional commands do not transform falsehood into truth. It does not authorize reckless breach of confidentiality or disrespect. It requires truthful and lawful courage ordered by conscience and responsibility.

FAITH AND VOCATION | The researcher as a witness to reality
A Christian vocation in research is not to force findings into a desired religious or political conclusion. It is to seek and report truth faithfully, recognize the limits of one's evidence, treat participants as neighbours rather than instruments, steward data carefully and resist incentives to distort. Christian conviction should make empirical work more honest, not less open to correction.

3.22 The Research Design Integrity Canvas: A Professional Toolkit

Before approving a study, a manager, supervisor, student or ethics reviewer can audit it through twelve connected questions.

Table 3.15. *The Research Design Integrity Canvas.*

Integrity question	What a strong answer contains	Warning sign
1. What problem is real?	Evidence of the practical condition and the unresolved uncertainty	Topic chosen only because data are easy to obtain
2. What exactly is being claimed?	Descriptive, associational, causal, interpretive or normative claim clearly distinguished	"Impact" used without a counterfactual design
3. What is the unit and population?	Unit of analysis, target population and inference boundary	Sample and population treated as interchangeable
4. What philosophical assumptions matter?	Only assumptions that materially affect evidence and interpretation	Paradigm named ceremonially with no design consequence
5. What prior evidence exists?	Transparent search and synthesis, competing findings and genuine gap	One-sided literature collection
6. Why this design?	Question-design fit and explicit comparison or interpretive logic	Method chosen from personal preference
7. Who or what will be sampled?	Frame, selection method and defensible size/adequacy rationale	Convenience sample presented as representative
8. How are concepts measured?	Operational definitions, instrument evidence and quality checks	Unvalidated items stand in for complex constructs
9. How will each question be analyzed?	Pre-specified mapping from evidence to answer	"Data will be analyzed using software X"
10. How are people and data protected?	Consent, risk, privacy, permissions, security and legal compliance	Ethics treated as a final signature
11. What could make the conclusion wrong?	Bias, confounding, missingness, rival explanations, limitations	Limitations listed generically after results
12. What will be disclosed and shared?	Transparent reporting, provenance, conflicts, reproducibility and correction plan	Sponsor controls the evidential narrative

Source/Note: Author's synthesis from sources cited in the surrounding text.

A design should not proceed because it scores well on most questions while failing catastrophically on one. A study with an invalid sampling frame, unprotected sensitive data or fabricated citations is not rescued by good formatting elsewhere.

PRACTICE TOOL | The legality-profitability-morality-evidence test
Applied business research often informs actions that can be legal, profitable and still poorly supported or morally wrong. Before recommending action, ask four separate questions: **Is it supported by evidence? Is it legally permissible? Is it economically viable? Is it morally defensible?** Do not collapse one test into another. Chapter 13 develops moral and governance reasoning in greater depth; Chapter 3 ensures the evidence side is worthy of the decision.

3.23 Common Misconceptions, Analytical Errors and Professional Pitfalls

Misconception 1: A large sample guarantees quality. Sample size cannot fix coverage bias, invalid measurement or a poor comparison. Ten thousand voluntary social-media responses can be less representative than a smaller probability sample.

Misconception 2: Quantitative research is objective and qualitative research is subjective. Both require judgment. Quantitative work embeds judgment in variable definition, model choice and data cleaning. Qualitative work can be disciplined, transparent and strongly constrained by evidence. The relevant question is how bias and interpretation are managed.

Misconception 3: Mixed methods is automatically stronger. Mixed methods adds value only when integration answers a question better than a single approach. Unintegrated methods multiply cost and inconsistency.

Misconception 4: Cronbach's alpha proves validity. Alpha concerns one form of internal consistency and is affected by item number and covariance. It does not prove unidimensionality, content validity or construct validity.

Misconception 5: A 5% margin of error always requires 384 respondents. That familiar figure comes from a particular simple-random-sample proportion calculation. Power, clustering, finite populations, rare outcomes and different estimands change requirements.

Misconception 6: Purposive sampling is inferior probability sampling. Purposive sampling is appropriate when the goal is theoretical or information-rich case selection. It becomes a problem when researchers make population prevalence claims that the sampling logic cannot support.

Misconception 7: Statistical significance establishes practical importance. Chapter 4 will show that statistical and practical significance are different. Research design should identify decision-relevant effect sizes before analysis.

Misconception 8: A cross-sectional association explains direction. Temporal ordering, confounding and reverse causation can make association ambiguous. Causal language should follow design strength.

Misconception 9: Ethics approval transfers responsibility to the committee. Researchers remain responsible for actual conduct, data security, amendments and emerging risks.

Misconception 10: Organizational permission equals participant consent. Gatekeepers authorize access; participants ordinarily decide whether their own participation is voluntary.

Misconception 11: Public data are ethically unrestricted. Public availability does not erase re-identification, contextual harm, contractual restrictions or platform terms.

Misconception 12: AI can do the literature review if it provides citations. Generated citations can be false or misapplied. Researchers must retrieve, read and verify sources and remain accountable for synthesis.

Misconception 13: Research gaps are discovered by adding a country name. Context matters only when it changes theoretical boundaries, evidence quality, mechanism or decision relevance.

Misconception 14: A method section should sound technical. The purpose of methodology is reproducible logic, not jargon. A reader should understand what was done and why.

Misconception 15: Limitations are weaknesses to hide. Every design has limits. Transparent boundaries make conclusions more credible because they show what the evidence does and does not support.

3.24 Chapter Synthesis: From Communicating Evidence to Designing It

This chapter has moved from truthful communication to truthful inquiry. Research methods are the architecture through which practical uncertainty becomes evidence. The chapter began by distinguishing practical problems, research problems and decision problems, then showed that methodology rests on assumptions about reality, knowledge, values and explanation. Philosophical literacy matters not because every student must adopt a permanent paradigm label, but because the kind of question asked should be coherent with the kind of evidence accepted.

The core discipline is alignment. A problem statement should lead to a purpose, objectives and questions. Questions should lead to a design. The design should define units, time horizons, comparisons, populations, samples and measures. Instruments should generate evidence capable of answering the question. Analysis should be planned before outcomes influence choices. Conclusions should be no stronger than the design allows.

Quantitative designs can estimate magnitudes and relationships and, under strong designs, causal effects. Qualitative designs can explain meaning, process and mechanism in context. Mixed methods can integrate these strengths when integration is explicit. Secondary and administrative data can provide scale and historical depth but must be interpreted through their data-generating processes. Sample size requires design-specific reasoning; qualitative adequacy depends on

information and analytic purpose rather than numerical folklore. Reliability, validity, trustworthiness and bias controls address different dimensions of evidence quality and should not be collapsed into a single “validity” claim.

Ethics and integrity surround the entire design. Consent, confidentiality, data protection, fair representation and refusal to fabricate are not optional afterthoughts. Rwanda's data-protection framework and wider African data-governance debates show that research increasingly operates across legal and technological borders. Generative AI makes verification, privacy and human accountability more important, not less.

The Nigeria YouWiN! and Indian management experiments illustrated how design transforms a plausible story into a stronger causal comparison while leaving external validity, implementation and moral judgment open to further analysis. Christian moral reasoning deepened these methodological duties by treating truth, human dignity, neighbour-love and stewardship as intrinsic obligations.

3.24.1 Hand-off to Chapter 4

Chapter 3 has taught the reader how defensible evidence is designed and generated. Chapter 4, **Data Analytics and Business Statistics: From Data Quality to Statistical Evidence**, asks what happens once quantitative data exist. It develops data quality, descriptive statistics, visualization, probability, sampling distributions, confidence intervals, hypothesis tests, regression and introductory forecasting. The sequence is deliberate. Statistical techniques cannot repair a sample that never represented the intended population, a variable that never measured the construct, or a design that never created a credible comparison. Chapter 4 therefore begins where Chapter 3 ends: with evidence whose provenance, purpose and limitations are already understood.

Key Terms

Abduction. Reasoning that seeks the most plausible explanation for an observed pattern and iteratively tests that explanation against evidence.

Accessible population. The portion of the target population that can realistically be reached through the available frame and procedures.

Anonymity. A condition in which participant identity is not known or reasonably linkable to the research data under the study design.

Axiology. Assumptions about values and their role in research, including the moral significance of problem choice and researcher conduct.

Bias. Systematic error that shifts estimates or interpretations away from the target of inference.

Case study. In-depth investigation of a bounded phenomenon in its real-world context, commonly using multiple evidence sources.

Causal inference. Reasoning about whether an exposure or intervention changed an outcome relative to a credible counterfactual.

Confidentiality. Protection of information and identities known to the research team from unauthorized disclosure.

Construct. An abstract concept represented through theoretically justified indicators or observations.

Construct validity. Evidence that a measure behaves in ways consistent with the theoretical construct it is intended to represent.

Counterfactual. The outcome that would have occurred for a unit under an alternative exposure or treatment condition.

Critical realism. A philosophical orientation that treats causal structures and mechanisms as real while recognizing that observation is partial and context-dependent.

Cross-sectional design. A design observing units at one time or during a limited period without following the same units over time.

Data-generating process. The institutional, technical and behavioral process through which observations enter a dataset.

Deduction. Reasoning from theoretical premises toward specific testable expectations.

Design effect. The ratio of sampling variance under a complex design to the variance under a simple random sample of the same nominal size; it helps quantify how clustering, stratification or weighting changes effective information.

External validity. The extent to which an inference can generalize across relevant populations, settings, treatments or periods.

Hypothesis. A testable expectation about a relationship, difference or parameter.

Induction. Reasoning that develops patterns, categories or propositions from observations.

Information power. A qualitative sample-adequacy principle under which richer and more specific information can justify fewer participants for a bounded aim.

Internal validity. The credibility of an inference, especially a causal inference, within the studied setting and design.

Interpretivism. An orientation emphasizing situated meaning and the interpretation of social reality by actors and researchers.

Longitudinal design. A design that observes change or sequence over time, often following the same units or cohorts.

Mixed methods. Intentional integration of qualitative and quantitative approaches to answer a research problem.

Non-probability sampling. Sampling in which units do not have known selection probabilities.

Ontology. Assumptions about the nature of reality and what kinds of phenomena exist.

Operationalization. Translation of a concept or construct into observable indicators, measures or procedures.

Paradigm. A broad constellation of assumptions about reality, knowledge, values and appropriate inquiry.

Population. The complete set of units about which a study intends to make a defined inference.

Postpositivism. An orientation that assumes an external reality while treating knowledge claims as fallible, theory-laden and open to empirical testing.

Pragmatism. A problem-centred orientation that evaluates methodological choices by their fit to inquiry and consequences while requiring coherent justification.

Preregistration. Time-stamped documentation of hypotheses, outcomes or analysis decisions before the relevant results are observed, used to distinguish planned confirmatory analyses from later exploration.

Probability sampling. Sampling in which eligible units have known non-zero selection probabilities.

Pseudonymization. Replacing direct identifiers with codes while retaining a separate means of linkage under controlled conditions.

Qualitative research. Systematic inquiry into meaning, experience, process, practice and context using evidence not primarily reduced to numerical variables.

Quantitative research. Systematic inquiry that operationalizes phenomena numerically to estimate magnitudes, differences, associations or effects.

Random assignment. Allocation to treatment conditions by a random process, supporting baseline comparability in expectation.

Reliability. Consistency of a measurement process under defined conditions.

Reproducibility. The extent to which reported results can be regenerated from the same data and documented analytic procedures; it is distinct from replication with new observations.

Research design. The logic linking a research question to evidence and inference through units, comparisons, time, sampling, measurement and analysis.

Research gap. A defensible unresolved uncertainty, contradiction, boundary, mechanism or methodological limitation in existing knowledge.

Research integrity. Honest, reliable, respectful and accountable conduct across problem formulation, data collection, analysis, collaboration, authorship, reporting, correction and stewardship of research resources.

Sampling frame. The operational list or selection mechanism from which a sample is drawn.

Saturation. A family of qualitative stopping concepts referring to diminishing addition of relevant codes, meanings or theoretical insight, defined according to the analytic tradition.

Secondary data. Data originally generated for another purpose or by another actor and subsequently analyzed for the current research question.

Statistical power. The probability, under specified assumptions, that a study will detect a defined effect when that effect exists; power depends on effect size, variability, sample design and decision thresholds.

Theoretical framework. An established theory or coherent set of theoretical propositions used to organize a study's concepts and expected relationships.

Trustworthiness. Qualitative quality criteria concerning credibility, transferability, dependability, confirmability and transparent interpretation.

Unit of analysis. The entity about which the research makes its primary claim.

Validity. The degree to which evidence and theory support the interpretation and use of a measure or inference for its intended purpose.

Concept-Check Questions

1. What is the difference between a practical problem, a research problem and a decision problem? Give one business example of each.
2. Why is “we used SPSS” not a methodology?
3. Distinguish ontology from epistemology and give one way each can affect research design.
4. Under what conditions is a hypothesis appropriate, and when would research questions without hypotheses be preferable?
5. Why is a literature review based on serial source summaries methodologically weak?
6. Distinguish a theoretical framework from a conceptual framework.
7. What makes a research gap more defensible than the claim that “no study has been conducted in Country X”?
8. Explain the difference between cross-sectional and longitudinal design.
9. Why can an observational association fail to support a causal claim?
10. What is the defining feature of mixed-methods research?
11. Distinguish the target population, accessible population and sampling frame.
12. Why is a sample of about 384 not a universal requirement?
13. What is the difference between internal consistency and construct validity?
14. Give three examples of bias that should be addressed during design rather than after analysis.
15. Distinguish confidentiality, anonymity and pseudonymization.

16. Why does organizational permission not normally replace individual consent?
17. What must a researcher understand about the data-generating process before using administrative data?
18. Identify four risks of using generative AI in research and one control for each.
19. What is preregistration intended to improve, and what can it not guarantee?
20. Why is explicit disclosure of limitations a strength rather than a sign of failed research?

Higher-Order Analytical and Critical-Thinking Questions

1. A ministry wants to know whether a new export-promotion programme "worked." Construct three different meanings of "worked" and show how each would require a different design and outcome set.
2. A multinational company surveys only employees who voluntarily attend a leadership seminar and finds very high engagement. Analyze the sampling and causal-inference problems before any statistical test is run.
3. Compare postpositivist and interpretive approaches to studying trust in mobile-money agents. What would each treat as evidence, and what important aspect might each miss?
4. A researcher finds a strong association between digital adoption and SME productivity. Develop at least four rival causal explanations and identify design changes that would help distinguish among them.
5. A systematic review follows PRISMA reporting perfectly but searches only one database and excludes all studies with null results that lack full-text open access. Evaluate the review's credibility.
6. Should researchers ever randomize access to a scarce beneficial intervention? Analyze the issue using causal identification, fairness, informed consent, opportunity cost and at least two ethical frameworks.
7. Under what circumstances could a smaller purposive sample provide better evidence than a larger probability sample? Be precise about the claim being made.
8. A research team translates a validated employee-autonomy scale into three languages using back-translation. Explain why linguistic equivalence alone may still be insufficient for cross-country comparison.
9. Evaluate the claim: "Public social-media data do not require ethics review because users already chose to publish them."
10. A sponsor offers to fund a study only if it can review and approve the final report before release. Design a research-governance agreement that protects legitimate confidentiality while preserving evidential independence.

Applied Case Problem: Diagnosing Export Delays on the Northern Corridor

Kivu Harvest Ltd., a hypothetical Rwandan food-processing SME, exports shelf-stable fruit products to customers in Kenya. During the last 12 months, management reports that some shipments arrived substantially later than expected. The sales team attributes the problem to customs clearance; the logistics team blames inconsistent transporter performance; the finance team argues that delayed foreign-exchange settlement causes dispatch postponement; and the managing director believes border delays are being exaggerated to justify poor internal planning.

The company has 420 shipment records, but fields have changed over the year. Only 280 records contain a complete border-clearance timestamp. Three transporters account for 85% of volume. Customers sometimes report arrival dates through WhatsApp messages rather than the ERP system. The firm wants an answer within ten weeks and can interview staff, transporters and a small number of customers. It has permission to use operational records but must not disclose customer names or commercially sensitive route prices.

Task. Design a defensible study without performing the statistical analysis. Your design should: define the practical and research problems; formulate no more than three research questions; state the unit(s) of analysis; identify which claims are descriptive, associational, causal or interpretive; specify a primary design and justify whether mixed methods are necessary; construct a sampling strategy for interviews; audit the completeness and data-generating process of shipment records; define at least five operational variables; propose a missing-data and validation strategy; specify the qualitative analysis logic; identify key biases and rival explanations; address confidentiality and organizational permissions; and state what the study could **not** conclude within ten weeks.

> **EVIDENCE CHECK | The case has no single correct method**
> A strong solution is one whose question, evidence and inference are aligned. A student may reasonably choose a primarily secondary-data design with targeted interviews, or a mixed-method explanatory design. The solution becomes weak when it promises causal attribution that the data and time horizon cannot support.

Ethical and Faith-Informed Dilemma: When the Sponsor Wants a Better Result

A consulting team evaluates an employee-training programme funded by a major corporate client. The pre-specified outcome is retention at six months. The treatment group shows no meaningful improvement. After seeing the result, the

sponsor asks the team to replace retention with “intent to stay,” a survey measure collected after training that shows a large difference. The sponsor argues that the new measure better reflects the programme's purpose and reminds the research lead that renewal of a substantial contract depends on a “constructive” report.

The lead researcher knows that intent to stay is a legitimate construct but was not the primary outcome and may not translate into actual retention. The sponsor has not asked the team to invent data. It has asked for a selective reframing that could cause readers to believe the programme achieved the outcome it was originally designed to improve.

Dilemma. What should the research lead do? Analyze the issue in terms of pre-specification, exploratory versus confirmatory analysis, sponsor rights, contract duties, truthfulness, stewardship, neighbour-love, professional courage and the legitimate possibility that the sponsor's alternative construct still deserves analysis. Draft the methodological disclosure that should accompany any secondary analysis.

Data and Evidence Exercise: Auditing a Business Dataset Before Analysis

Obtain an openly available business or trade dataset, such as a World Bank Enterprise Survey dataset or a national statistical dataset for which use is permitted. Before calculating substantive results, produce a two-page **data provenance and fitness audit** containing:

1. The dataset's producer, original purpose, collection period and version.
2. The target population and explicit universe of inference.
3. The unit of observation and unit of analysis.
4. The sampling or enumeration method, including strata or clusters where applicable.
5. Five variables central to a plausible business question, with their operational definitions and units.
6. A missingness table and explanation of whether missing codes have multiple meanings.
7. Evidence of classification or definition changes that could affect comparison over time.
8. At least three plausible sources of bias that analysis alone cannot fully remove.
9. The legal, licensing and ethical conditions governing use.
10. One research question the dataset can credibly answer and one tempting question it cannot answer without additional evidence.

The exercise is complete only when the student can explain the data-generating process. Do not begin regression or hypothesis testing until this audit is defensible.

References

African Union. (2022). *AU data policy framework*. https://aucapps.au.int/en/documents/20220728/au-data-policy-framework

All European Academies. (2023). *The European code of conduct for research integrity (Revised edition 2023).* https://doi.org/10.26356/ECOC

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does management matter? Evidence from India. *The Quarterly Journal of Economics, 128*(1), 1-51. https://doi.org/10.1093/qje/qjs044

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*. Heinemann.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Houghton Mifflin.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). Wiley.

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334. https://doi.org/10.1007/BF02310555

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. SAGE.

Data Protection and Privacy Office. (n.d.). *Services*. Retrieved August 27, 2026, from https://dpo.gov.rw/services

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532-550. https://doi.org/10.5465/amr.1989.4308385

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. *Health Services Research, 48*(6 Pt 2), 2134-2156. https://doi.org/10.1111/1475-6773.12117

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods, 16*(1), 15-31. https://doi.org/10.1177/1094428112452151

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). SAGE.

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: How many interviews are enough? *Qualitative Health Research, 27*(4), 591-608. https://doi.org/10.1177/1049732316665344

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*(1), 104-121. https://doi.org/10.1177/109442819800100106

Horn, L., Alba, S., Gopalakrishna, G., Kleinert, S., Kombe, F., Lavery, J. V., & Visagie, R. G. (2023). The Cape Town Statement on fairness, equity and diversity in research. *Nature, 615*(7954), 790-793. https://doi.org/10.1038/d41586-023-00855-y

Kish, L. (1965). *Survey sampling*. Wiley.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE.

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research, 26*(13), 1753-1760. https://doi.org/10.1177/1049732315617444

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). SAGE.

McKenzie, D. (2017). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. *American Economic Review, 107*(8), 2278-2307. https://doi.org/10.1257/aer.20151404

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*(1), 48-76. https://doi.org/10.1177/1558689806292462

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour, 1*, 0021. https://doi.org/10.1038/s41562-016-0021

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. U.S. Department of Health and Human Services. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences, 115*(11), 2600-2606. https://doi.org/10.1073/pnas.1708274114

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879

Republic of Rwanda. (2021). *Law No. 058/2021 of 13/10/2021 relating to the protection of personal data and privacy*. Official Gazette of the Republic of Rwanda. https://www.risa.gov.rw/fileadmin/user_upload/RISA/Publications/2.Laws/Law_relating_to_the_protection_of_personal_data_and_privacy.pdf

Resnik, D. B., & Hosseini, M. (2026). Hallucinated citations produced by generative artificial intelligence may constitute research misconduct when citations function as data in scholarly papers. *Accountability in Research.* Advance online publication. https://doi.org/10.1080/08989621.2026.2645390

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2023). *Research methods for business students* (9th ed.). Pearson.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349-357. https://doi.org/10.1093/intqhc/mzm042

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207-222. https://doi.org/10.1111/1467-8551.00375

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLOS Medicine, 4*(10), e296. https://doi.org/10.1371/journal.pmed.0040296

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*, 160018. https://doi.org/10.1038/sdata.2016.18

World Bank Enterprise Surveys. (n.d.). *Methodology*. World Bank Group. Retrieved August 27, 2026, from https://www.enterprisesurveys.org/en/methodology

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.

Annotated Further Reading

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches (6th ed.).* A clear cross-paradigm guide to qualitative, quantitative and mixed-methods design, especially useful for connecting philosophical assumptions, research questions and design choices.

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2023). *Research methods for business students (9th ed.).* A comprehensive business-focused methods guide covering philosophy, access, sampling, data collection and analysis; particularly useful as a procedural companion for applied business research.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* A foundational guide to causal design, validity threats and generalization, essential for understanding why causal inference depends first on research architecture.

Yin, R. K. (2018). *Case study research and applications: Design and methods (6th ed.).* A rigorous guide to defining cases, selecting evidence and building analytic logic, valuable for avoiding the common misuse of the case-study label.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). *Seeking qualitative rigor in inductive research: Notes on the Gioia methodology.* An influential account of transparent inductive organizational analysis and theory development, best used as one methodological approach rather than a universal qualitative template.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). *Achieving integration in mixed methods designs: Principles and practices.* A practical explanation of integration in mixed methods, useful for designing studies in which qualitative and quantitative components genuinely inform one another.

McKenzie, D. (2017). *Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition.* A strong African field experiment illustrating random assignment, longitudinal follow-up, entrepreneurship outcomes and the distinction between causal effects and universal policy claims.

Snyder, H. (2019). *Literature review as a research methodology: An overview and guidelines.* A concise guide to literature review as a research method, especially useful for designing transparent, defensible evidence syntheses in business research.

Horn, L., Alba, S., Gopalakrishna, G., Kleinert, S., Kombe, F., Lavery, J. V., & Visagie, R. G. (2023). *The Cape Town Statement on fairness, equity and diversity in research.* An Africa-rooted integrity statement on fair agenda setting, authorship, data governance, capacity and collaboration in cross-border research partnerships.

All European Academies. (2023). *The European code of conduct for research integrity (Revised edition 2023).* A concise integrity framework organized around reliability, honesty, respect and accountability, useful for research governance alongside national law and disciplinary standards.

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research.* A policy-oriented guide to responsible generative-AI use, privacy and human agency, to be read alongside current institutional and publisher rules.