

Qualitative Research Methods

Foundations, Design, Analysis and Tools

Research Methods Seminar

University of Nairobi, Kenya

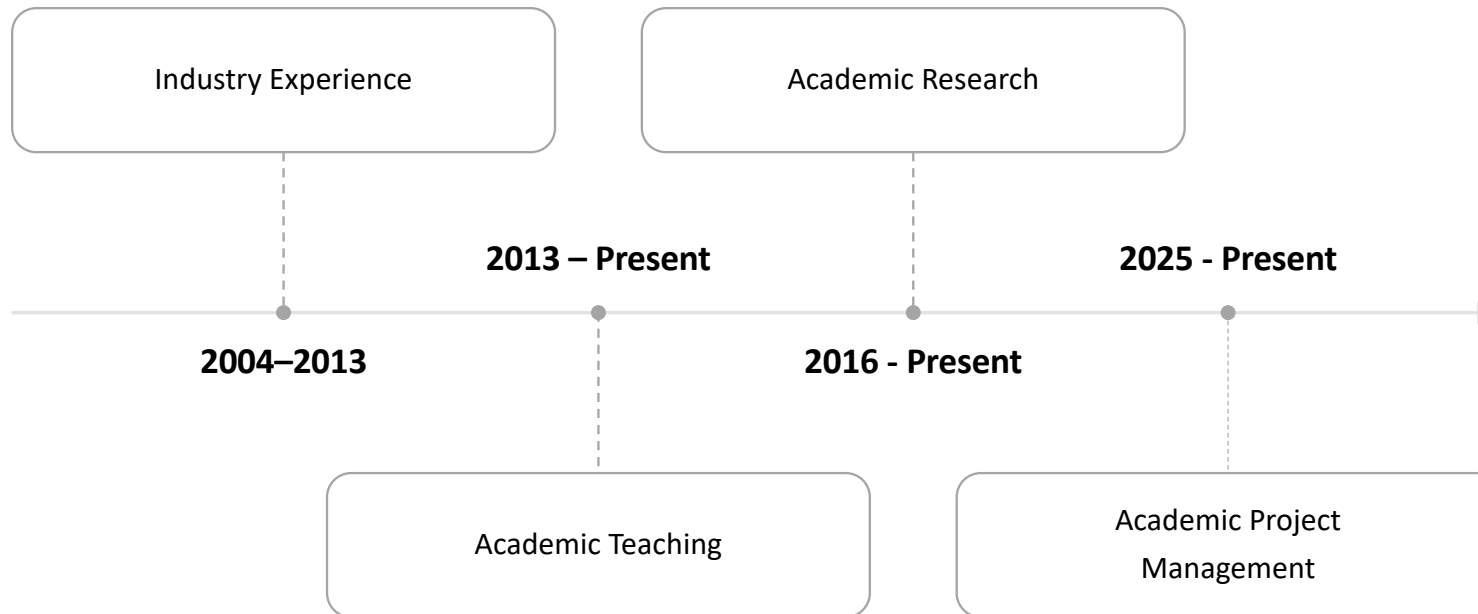
Dr. Simon Peter Nadeem

Senior Program Manager,

LEARN Logistics by Kuehne Foundation

Dr. Simon Peter Nadeem

Dr. Simon Peter Nadeem



Dr Simon Peter Nadeem
PhD, FHEA, MBA, E-MBA



Google Scholar profile

<https://scholar.google.co.uk/citations?user=vrCvUt8AAAAJ&hl=en>

Today's Programme

Time	Session	What you will be able to do afterwards
09:00 – 09:50	1. Foundations of Qualitative Research	Explain what qualitative research is, the assumptions behind it, and when it is the right choice
10:00 – 10:50	2. Research Design and Data Collection	Turn a topic into a researchable question, match it to a design, justify your sample and choose methods
10:50 – 11:10	Coffee / Tea break	
11:10 – 11:50	3. Qualitative Analysis and Rigour	Code data, build themes, and defend the trustworthiness of what you claim
12:00 – 13:00	4. Tools for Qualitative Research and the Role of AI	Choose software on methodological grounds and use AI without compromising rigour, ethics or the law

Ground rules: questions at any point • no such thing as a naive methods question • bring your own project to every activity

Learning Outcomes

By the end of this seminar, you should be able to...

1	Position	Distinguish qualitative from quantitative inquiry, and locate both within a philosophical position (ontology, epistemology, methodology)
2	Design	Formulate a qualitative research question and select a methodology that genuinely answers it
3	Sample	Justify a purposive sample and a sample size using saturation and information power — not convenience
4	Collect	Select and design instruments for interviews, focus groups, observation and documents
5	Analyse	Code data and construct themes using a named, defensible analytic approach
6	Defend	Apply trustworthiness criteria and reporting standards so a reviewer can audit your work
7	Tool up	Choose qualitative software on methodological grounds, and use AI critically, ethically and legally

SESSION 1 • 09:00 – 09:50

Foundations of Qualitative Research

- What research is, and what makes it scientific
- Deduction, induction, abduction
- What qualitative research is — and what it is not
- Philosophical grounding: ontology, epistemology, methodology
- Qualitative and quantitative compared
- When qualitative research is the right choice

What is Research?

“Research is finding things out.”

...is that enough?

We acquire knowledge in many ways

- Asking a friend, a colleague, a teacher
- Searching the internet, books, reports
- Watching television, reading newspapers and social media

But how do we know we are establishing what is true, and not simply circling back to what we already believed?

Research is systematic inquiry that makes its own reasoning inspectable — so that someone else can judge, challenge and build on it.

What Makes Research Scientific?

Four features that separate research from opinion

Grounded in what is known

- Determines what still needs to be found out by reviewing existing scholarly knowledge
- Usually by systematic engagement with peer-reviewed literature

Theoretically framed

- Formulates the problem by reference to particular theories and concepts
- A topic is not yet a research problem

Methodologically rigorous

- Uses methods capable of producing surprising findings
- Not designed merely to confirm what we already expect

Contributes knowledge

- Often interested in knowledge for its own sake, not only immediate practical use
- Contribution must be stated, not assumed

A useful test for your own project: could your method produce a finding you would not like?

Framing adapted from Silverman, D. (2017) Doing Qualitative Research, 5th edn. London: Sage; and Saunders, M.N.K., Lewis, P. & Thornhill, A. (2023) Research Methods for Business Students, 9th edn. Harlow: Pearson.

Scientific Reasoning: Three Routes to Theory

Qualitative research is not automatically inductive — be explicit about which route you are taking

Deduction

Theory → Data

From general premises to a specific conclusion

- All birds lay eggs; penguins are birds; therefore penguins lay eggs
- In research: test a proposition drawn from existing theory
- Risk: you only ever see what your framework allows

Induction

Data → Theory

From specific observations to a general claim

- Two friends choose pizza; you infer pizza is good and order it too
- In research: build concepts and theory up from the data
- Risk: over-generalising from a thin base

Abduction

Puzzle → Best explanation

From partial evidence to the most plausible explanation

- Treating traces at a scene as clues to what happened
- In research: move back and forth between data and theory
- Most honest description of how good qualitative work actually proceeds

In your methods chapter, name the logic you used — and show where it changed. Abductive is a legitimate and increasingly common answer.

What is Qualitative Research?

Qualitative Research: Working Definitions

Three complementary framings, not three competing ones

“Qualitative research is concerned with meaning — how people make sense of their world and their experience of it — and works primarily with words, images and other non-numeric material.”

A serviceable working definition for this seminar

It asks different questions

- What is going on here? How does this work? What does this mean to the people involved?
- Rather than: how many, how often, how much, how strongly correlated

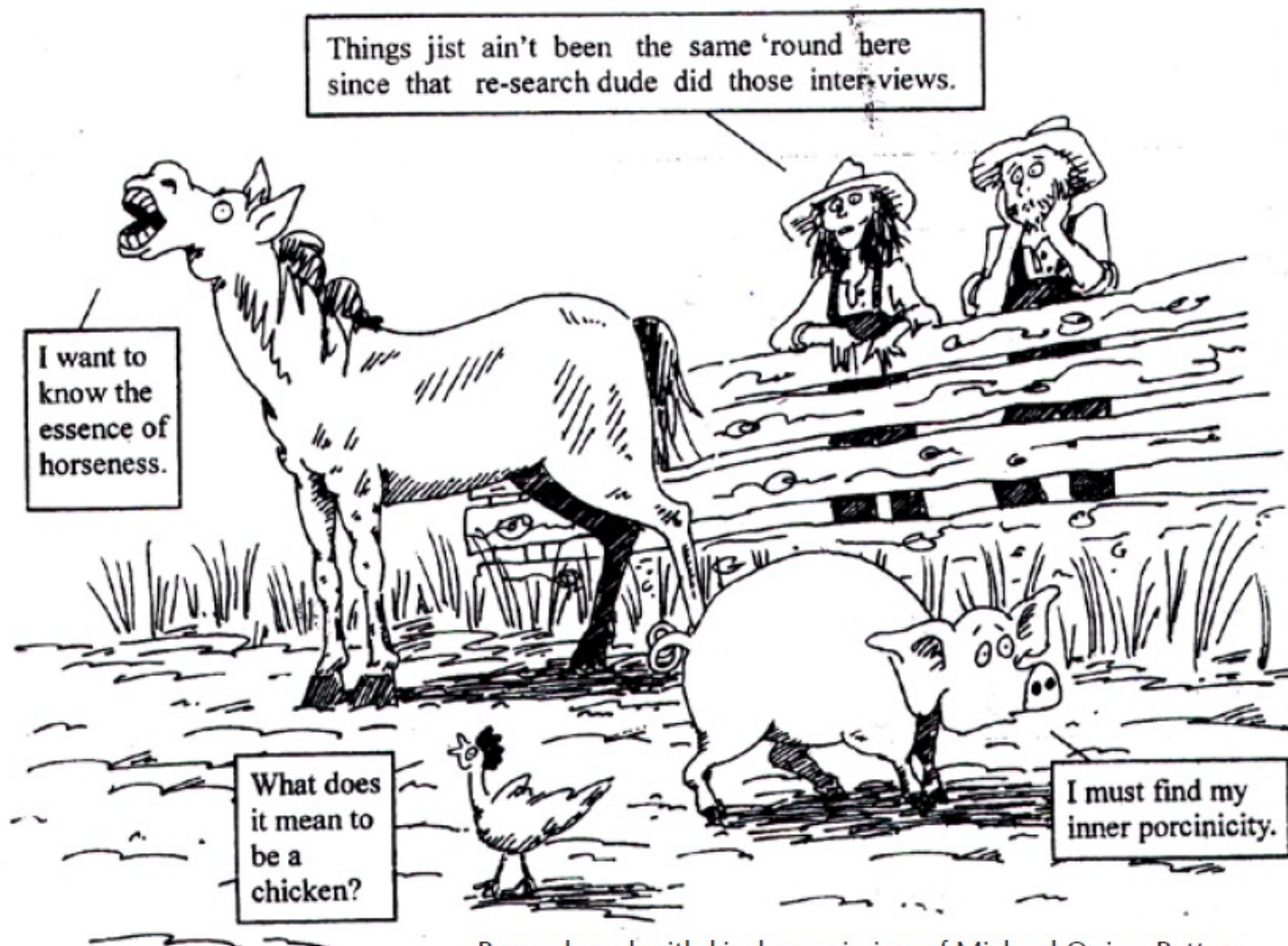
It treats context as data, not noise

- The setting, the relationships and the researcher's presence are part of what is being studied

It is interpretive

- Findings are produced through the researcher's engagement with the data — which is precisely why the process must be made transparent

**Qualitative does not mean 'without numbers', 'small-scale', or 'soft'.
It means the analytic object is meaning rather than magnitude.**



© 2002 Michael Quinn Patton and Michael Cochran

Reproduced with kind permission of Michael Quinn Patton

A Warning Before We Go Further

On taking familiar quotations on trust

*“Not everything that can be counted counts,
and not everything that counts can be counted.”*

attributed to Albert Einstein

- The line is a fair summary of why qualitative research exists.
- The attribution to Einstein is repeated everywhere and is not supported by any primary source. It is usually traced instead to twentieth-century sociological writing.

The methodological point

- If a claim this famous circulates unchecked for decades, assume the same can happen to a citation in your literature review.
- Verify every quotation and every reference against the original. Today's fourth session will show you how quickly AI tools make this problem worse.

Where Qualitative Research Comes From

A set of traditions rather than a single method — which is why the vocabulary varies

Anthropology

- Ethnography, fieldwork, thick description of culture and practice

Sociology

- Grounded theory, symbolic interactionism, the study of social order and interaction

Psychology

- Phenomenology and IPA — the structure of lived experience

Linguistics & Communication

- Discourse and conversation analysis, semiotics, narrative

Education & Health

- Case study, framework and content analysis oriented to practice and policy

Management & Logistics

- Case study research, process research, practice-based studies of organisations and supply chains

Practical consequence: the same word can mean different things in different traditions. Define your terms and cite the tradition you are working in.

Three Questions Underneath Every Design

You answer these whether or not you write them down — better to write them down

1	Ontology	What do I think reality is like? Is there one social reality to be discovered, or multiple realities constructed by people?
2	Epistemology	What counts as acceptable knowledge, and what is my relationship to what I study? Can I be a detached observer, or am I part of it?
3	Axiology	What role do my values play? Are they contamination to be controlled, or a resource to be declared and used reflexively?
4	Methodology	Given the three answers above, what strategy of inquiry follows? Methodology is the justification; method is the technique.

Alignment is the marking criterion: a positivist ontology with an interpretive analysis and a convenience sample is not a design, it is a collision.

Five Research Philosophies

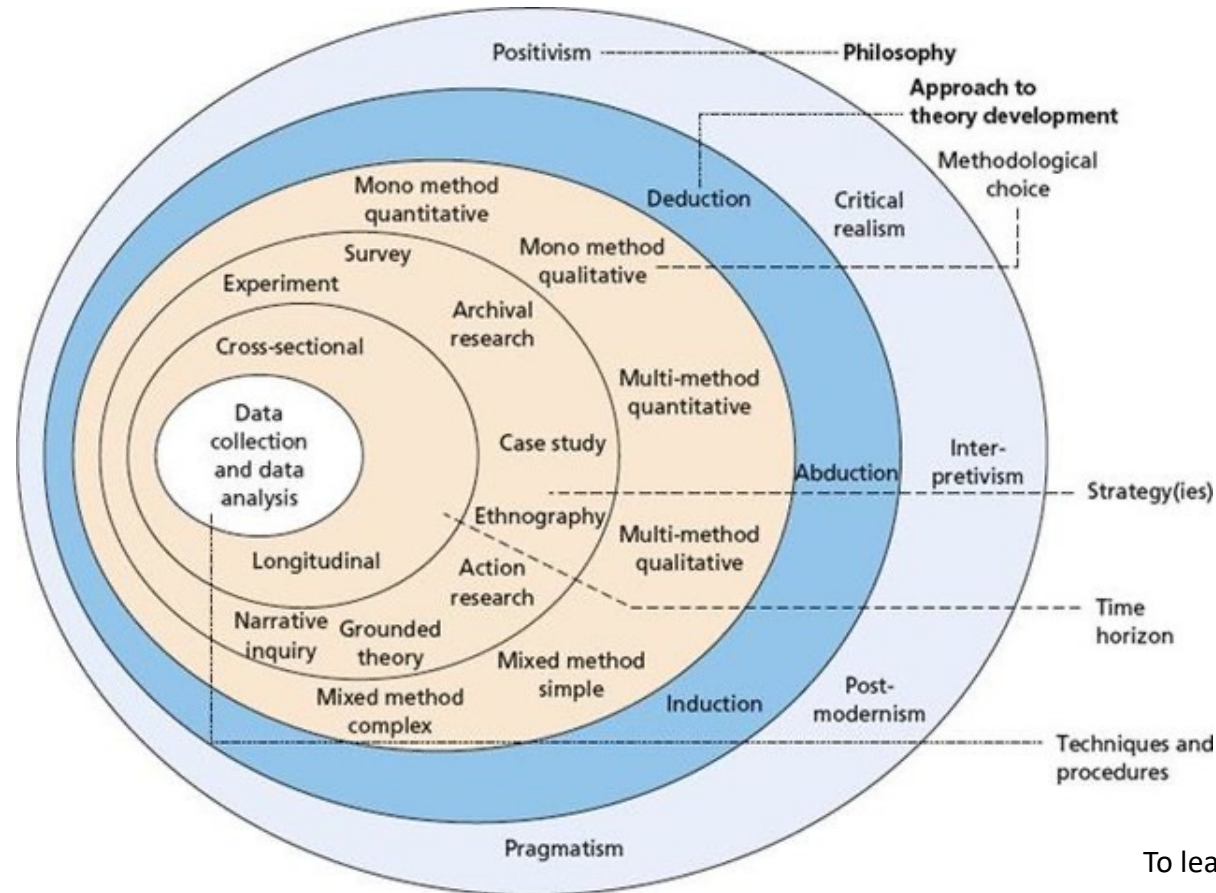
Where does your view of the world sit?

Philosophy	Core assumption	Typical qualitative use
Positivism	An objective reality exists independently of us and can be measured; the researcher stays detached	Rare as a qualitative stance; sometimes used for structured content analysis with coder reliability
Critical realism	A real world exists, but we only ever access it through our interpretations, which are fallible and historically situated	Case studies seeking underlying mechanisms behind observed events
Interpretivism	Social reality is made and remade through meanings people attach to it; understanding, not explanation, is the goal	The default home of most qualitative work: ethnography, phenomenology, thematic analysis
Postmodernism / poststructuralism	Claims to objective knowledge conceal power; language constitutes rather than describes reality	Discourse analysis, deconstruction, studies of silence and marginalisation
Pragmatism	Start from the problem; the value of knowledge lies in what it lets us do	Mixed methods designs, applied and policy-oriented research

Adapted from Saunders, M.N.K., Lewis, P. & Thornhill, A. (2023) Research Methods for Business Students, 9th edn. Harlow: Pearson. The HARP reflexive tool (Bristow & Saunders, 2014) in that text helps you locate your own position.

Design as a Series of Nested Choices

The 'research onion' — you must be able to justify every layer, from the outside in



To learn more, visit the following website:

<https://gradcoach.com/saunders-research-onion/>

Saunders, Lewis & Thornhill (2023) Research Methods for Business Students, 9th edn. Harlow: Pearson.

Dr. Simon Peter Nadeem

Qualitative and Quantitative Compared

Two different jobs — read the table as a contrast of purposes, not a ranking

	Qualitative research	Quantitative research
Typical purpose	Theory building; understanding and interpretation	Theory testing; description and explanation of variance
Question form	What? How? Why in this case?	How many? How often? How strongly related?
Researcher role	High involvement; the researcher is part of the setting	Limited, controlled involvement; distance is designed in
Sampling	Non-probability, purposive; small and deliberately chosen	Probability where possible; large enough for inference
Data	Words, images, artefacts, interaction; reduced to codes and themes	Numbers, or observations reduced to numeric codes
Analysis	Human interpretation, systematically evidenced; software assists, it does not decide	Statistical, computerised; emphasis on counts and tests
Claim to generality	Transferability: readers judge fit to their context, aided by rich description	Statistical generalisation to a defined population

Adapted from Silverman, D. (2017) Doing Qualitative Research, 5th edn, and Silverman, D. (2019) Interpreting Qualitative Data, 6th edn. London: Sage.

Quantitative Methods



Qualitative Methods



Four Things Qualitative Research Is Not

These misconceptions cost marks in vivas and reviews

Not "quantitative without the maths"

- **Myth: it is the easier option for people uncomfortable with statistics.**
- It demands a different rigour — analytic transparency, reflexivity and an audit trail — and it is usually more labour-intensive per participant.

Not merely descriptive

- **Myth: you summarise what people said.**
- Summarising is the first step, not the finding. The work is explaining why the pattern is there and what it means theoretically.

Not unable to generalise

- **Myth: $n=15$ means the results mean nothing.**
- Qualitative work generalises to theory and to transferable insight, not to a population. State which claim you are making.

Not inferior to quantitative work

- **Myth: it is a preliminary stage before the real study.**
- Some quantitative research is weak; so is some qualitative research. Quality is a property of execution, not of paradigm.

Should I Use Qualitative Research?

Six questions to ask before you commit

1	Purpose	What exactly am I trying to find out? Different questions require different methods — the question governs the choice, never the reverse.
2	Focus	Do I want depth on a phenomenon in context, or systematic comparison and accounting for variance?
3	Field	How have other researchers dealt with this topic, and how far do I want to align with that literature?
4	Feasibility	How long will this take? Can I get access to the case I want to study in depth? What data already exists?
5	Payoff	What will I learn with this approach that I could not learn with the other? What is the knowledge payoff?
6	Fit to me	Am I committed to a research model that implies a methodology? Do I have the skills, or time to build them?

Adapted from Punch, K. (1998) Introduction to Social Research, as presented in Silverman, D. (2017) Doing Qualitative Research, 5th edn. London: Sage.

A Published Example: Seeing What Interviews Miss

Session 1 case — why context is data

PUBLISHED EXAMPLE

Jepkosgei, J., English, M., Adam, M.B. & Nzinga, J. (2022) '*Understanding intra- and interprofessional team and teamwork processes by exploring facility-based neonatal care in Kenyan hospitals*', BMC Health Services Research, 22, Article 636. doi:10.1186/s12913-022-08039-6 (open access)

Design & data: Ethnographic approach across three hospitals in Nairobi and Murang'a Counties: 250 hours of non-participant observation over nine months, plus 17 in-depth interviews with 20 clinicians and nurses. Thematic content analysis, coded independently and then reconciled to an agreed framework.

Why it is a useful example: The observation log is methodologically heavier than the interview set. From watching wards, the team identified three kinds of team — a core team, an emergency team, and temporary ad-hoc teams — categories that no interview question had asked about. It also models an auditable coding process: independent coding, then team consensus, then a final framework.

Talking point: if the study had only interviewed staff, would the ad-hoc teams ever have been named?

Ask of your own project: what would I only ever find out by being there?

Activity (in Pairs)

Session 1 activity

1 — Name it

- Describe to your partner one thing about your field you do not understand well enough.
- Not a gap in the literature - a genuine puzzle about how something works or what it means to people.

2 — Test it

- Could this be answered by counting? If yes, it is probably a quantitative question.
- If the answer depends on meaning, process or context, it is a candidate for qualitative inquiry.

3 — Locate it

- Which philosophical position does your puzzle assume - is there one truth to uncover, or several accounts to understand?
- Keep this note. You will use it in next part.

We will take two or three examples from the room before the break — volunteers welcome, and disagreement is more useful than consensus.

Session 1 — What to Carry Forward

Four points

1

Research is inspectable inquiry

What distinguishes research from opinion is not certainty but the fact that its reasoning can be examined by others.

2

Qualitative work studies meaning

The analytic object is what things mean and how they work in context — not how much or how many.

3

Assumptions are unavoidable

Ontology, epistemology and axiology shape your design whether or not you declare them. Declare them.

4

The question governs the method

Choose qualitative because your question demands it — never because it looks easier or the data is nearer.

SESSION 2 • 10:00 – 10:50

Research Design and Data Collection

- From topic to a researchable qualitative question
- Five methodologies and what each commits you to
- Purposive sampling, saturation and information power
- Interviews, focus groups, observation, documents
- Designing the instrument — and what to avoid
- Ethics, access and the Kenyan regulatory context

From Topic to Research Question

Most weak qualitative projects are weak here, not in the analysis

Topic

Broad area of interest

“Warehouse safety”

Puzzle

Something specific you do not understand

“Why do trained staff bypass safety procedures they helped design?”

Question

Answerable, bounded, and phrased for meaning or process

“How do warehouse operatives account for departures from agreed safety procedures?”

Sub-questions

Two to four, each answerable with identifiable data

What situations prompt departures? How are they justified to peers? How do supervisors respond?

Diagnostic: if your question can be answered “yes”, “no” or with a number, it is not yet a qualitative research question.

Weak and Strong Qualitative Questions

Same territory, very different research

Weak

- X “Is supplier trust important?” — answerable yes/no; nothing to analyse
- X “What are the factors affecting procurement performance?” — a list, and implicitly a survey
- X “Why do SMEs fail to adopt digital tools?” — presumes failure, and presumes the cause is in the SME
- X “How can we improve last-mile delivery?” — a consultancy brief, not a research question
- X “What is the impact of X on Y?” — a causal-effect question that qualitative design cannot answer as posed

Stronger

- ✓ “How do buyers and suppliers construct and repair trust after a disruption?”
- ✓ “How do procurement officers reconcile competing pressures when choosing suppliers?”
- ✓ “How do SME owners account for their decisions about digital tools, and what do those accounts reveal about perceived risk?”
- ✓ “How do delivery riders navigate the gap between routing systems and street-level realities?”
- ✓ “How do participants themselves explain the relationship between X and Y?”

Note the verbs: how, in what ways, how do participants account for. They commit you to meaning and process.

Alignment: The Thing Examiners Actually Check

Read each row across — a break anywhere invalidates the row

If your question is about...	The methodology that fits	Data you will need	What you can then claim
The culture and practices of a setting	Ethnography	Prolonged observation, field notes, informal conversation, documents	A situated account of how things actually work
A social process nobody has yet theorised	Grounded theory	Iterative interviews with theoretical sampling, memos	A theory of that process, grounded in data
The texture of lived experience	Phenomenology / IPA	A small number of deep, unstructured or lightly structured interviews	The essential structure and meaning of an experience
How and why something happened in a bounded instance	Case study	Multiple sources within the case: interviews, documents, observation, records	An explanation of the case; propositions transferable to similar cases
How people story their lives and events	Narrative inquiry	Extended narrative-eliciting interviews, sometimes repeated over time	How meaning is constructed and organised over time

You do not need an exotic methodology. You need one you can name, justify and execute properly.

What Each Methodology Commits You To

The obligations, not just the labels

Ethnography

- **Commits you to: time in the field.**
- Weeks or months, not visits. Field notes are primary data.
- Fails when: it becomes a few interviews on site.

Grounded theory

- **Commits you to: analysis during collection.**
- Theoretical sampling, constant comparison, memos, saturation of categories.
- Fails when: all data is gathered first, then coded.

Phenomenology / IPA

- **Commits you to: bracketing and depth.**
- Homogeneous sample, few participants, focus on experience as experienced.
- Fails when: it becomes an opinion survey.

Case study

- **Commits you to: defining the case and its boundary.**
- Multiple data sources triangulated within the case; explicit case-selection logic.
- Fails when: “case study” just means “one organisation”.

Narrative inquiry

- **Commits you to: keeping stories whole.**
- Analyse structure and sequence, not just extracted fragments.
- Fails when: narratives are chopped into decontextualised codes.

Content / framework analysis

- **Commits you to: an explicit, documented coding frame.**
- Often a matrix by case and category; suits policy and multi-site work.
- Fails when: the frame is never reported.

On case study design: Eisenhardt, K.M. (1989) 'Building theories from case study research', Academy of Management Review, 14(4), 532–550; Yin, R.K. (2018) Case Study Research and Applications: Design and Methods, 6th edn. Thousand Oaks, CA: Sage. On grounded theory: Glaser, B.G. & Strauss, A.L. (1967) The Discovery of Grounded Theory. Chicago: Aldine.

A Published Example: Grounded Theory Done Properly

Case — what makes a study grounded theory rather than “interviews plus themes”

PUBLISHED EXAMPLE

Hennegan, J., Schwab, K.J., Exum, N.G., Kibira, S.P.S., Makumbi, F.E. & Bukenya, J. (2020) *“I do what a woman should do”: a grounded theory study of women’s menstrual experiences at work in Mukono District, Uganda*, *BMJ Global Health*, 5(11), e003433. doi:10.1136/bmjgh-2020-003433 (open access)

Design & data: Grounded theory with 35 in-depth interviews across three deliberately different occupational groups (21 market workers, 7 teachers, 7 health-facility workers). Collaborative analysis sessions ran throughout data collection, supported by participant, workplace and category memos.

Why it is a useful example: It shows the three things that actually distinguish grounded theory: analysis interleaved with fieldwork, memo-writing as a recorded analytic device, and the derivation of a single core category — “being a responsible woman” — that organises everything else. The purposive variation across three occupations is a visible sampling logic you can copy.

The whole theory is compressed into one participant’s phrase, which became the title.

Check: does your proposed grounded theory study have a plan for analysing between interviews? If not, it is thematic analysis — which is fine, but call it that.

A Published Example: One Case, Defensible Claim

Case — critical-case selection in a supply chain setting

PUBLISHED EXAMPLE

Botes, A., Niemann, W. & Kotzé, T. (2017) *'Buyer–supplier collaboration and supply chain resilience: a case study in the petrochemical industry'*, South African Journal of Industrial Engineering, 28(4), 183–199. doi:10.7166/28-4-1736 (open access)

Design & data: A single case study in the South African petrochemical industry, selected explicitly as a “critical case”. Semi-structured interviews, analysed against Kraljic's procurement matrix and the supply chain resilience literature.

Why it is a useful example: It answers the question students always ask — how can one case be enough? The answer is the selection logic: a critical case is chosen because if the relationship does not hold here, it is unlikely to hold anywhere. It also shows theory-informed rather than atheoretical qualitative work.

Headline finding: collaboration did not directly produce resilience, but enabled its antecedents — a qualitative study refining a causal story that quantitative work had assumed.

Case-selection logics worth naming: typical, critical, extreme/deviant, revelatory, longitudinal, and convenient (be honest if it is the last one).

Sampling in Qualitative Research

Not a smaller version of probability sampling — a different logic entirely

- Sampling means selecting units (people, organisations, documents, departments, moments in time) with direct reference to the research question.
- There are two levels: first you select the case or cases, then you select units within the case.
- It is deliberately non-random. The aim is to discover categories and their properties, not to represent a population.

Purposive

- Selected against explicit criteria derived from the research question. The default logic; all the others are variants.

Theoretical

- Who to approach next is decided by what the emerging analysis still needs. Requires concurrent analysis (Glaser & Strauss, 1967).

Maximum variation

- Deliberately diverse cases to map the range of a phenomenon and test the limits of a pattern.

Snowball

- Initial contacts refer others. Practical for hidden or hard-to-reach populations; brings network bias you must report.

Deviant / extreme case

- The outlier chosen on purpose, because it exposes what the normal case conceals.

Convenience

- Whoever is available. Legitimate only when acknowledged as a limitation — never dressed up as purposive.

Theoretical sampling and theoretical saturation: Glaser, B.G. & Strauss, A.L. (1967) The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago: Aldine.

“How Many Interviews Do I Need?”

The honest answer is: enough to answer your question — and here is how to argue it

Saturation

The traditional answer

- The point at which additional data yields no new codes, categories or properties.
- **Evidence: Guest, Bunce & Johnson (2006) analysed 60 in-depth interviews in West Africa and found saturation occurred within the first 12 interviews, with the basic elements of meta-themes present as early as 6.**
- Caution: saturation is a claim you must evidence, not assert. Show when new codes stopped appearing.

Information power

The better modern framing

- The more information your sample holds, the fewer participants you need. Five dimensions to argue:
- **Aim — narrow aims need fewer participants than broad ones**
- **Sample specificity — highly specific participants carry more information**
- **Theory — established theory reduces the data needed**
- **Dialogue quality — strong interviews carry more than weak ones**
- **Analysis strategy — in-depth case analysis needs fewer than cross-case comparison**

Never write “saturation was reached” without evidence. Write what you did, at what point new codes stopped emerging, and why that was sufficient given your aim.

Guest, G., Bunce, A. & Johnson, L. (2006) 'How many interviews are enough? An experiment with data saturation and variability', *Field Methods*, 18(1), 59–82. doi:10.1177/1525822X05279903 • 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. doi:10.1177/1049732315617444

Ways of Generating Qualitative Data

Each one answers a different kind of question — choose, do not default

Method	Best for	Main limitation
In-depth / semi-structured interviews	Accounts, reasoning, meaning, retrospective sense-making	You get what people say they do, mediated by what they want you to think
Focus groups	Shared norms, disagreement, how views are formed and defended in company	Group effects: dominant voices, self-censorship, socially desirable answers
Observation	Practice as it happens, routines, the gap between account and action	Reactivity; time-intensive; covert observation rarely receives ethical approval
Documents and records	Institutional reasoning, change over time, the unspoken official version	Produced for other purposes; silences are hard to read
Diaries, logs and journals	Experience close to the moment; sensitive or hard-to-observe topics	Attrition, uneven completion, participant burden
Photo-elicitation and participatory methods	Prompting recall, surfacing what words do not reach, shifting power to participants	Consent complexity, especially with images of third parties

Definitions of diaries, logs and journals after Hedlund et al. (1989) and Chabon & Lee-Wilkerson (2006); photo-elicitation after Harper, D. (2002); observation after Ghauri & Gronhaug (2002), as presented in the seminar handbook.

Designing the Interview

Question types, and the mistakes that cost you data

Types of question

Kvale (1996)

- **Introducing** — “Tell me about...”
- **Follow-up** — pursuing what was just said
- **Probing** — “Can you say more about that?”
- **Specifying** — “What happened next?”
- **Direct and indirect** — “What do most people think about...?”
- **Structuring** — “Let's move on to...”
- **Silence** — the most underused instrument in the set
- **Interpreting** — “Do you mean that...?”

What to avoid

Common design failures

- Ambiguous quantifiers: often, regularly, frequently
- Two questions in one — you will get one answer
- Leading questions that hint at a preferred response
- Very general questions with no frame of reference
- Negatives, and especially double negatives
- Jargon, technical terms and acronyms
- Over-reliance on yes/no questions
- Questions that depend heavily on distant memory

Always pilot. Piloting checks flow, exposes vague questions, and lets you drop items that produce uniform answers — and people who help you pilot should not appear in your final sample.

A Published Example: Focus Groups With a Job to Do

Case — qualitative work as the explanatory arm of a trial

PUBLISHED EXAMPLE

Kedenge, S.V., Kangwana, B.P., Waweru, E.W., Nyandigisi, A.J., Pandit, J., Brooker, S.J., Snow, R.W. & Goodman, C.A. (2013) *'Understanding the impact of subsidizing artemisinin-based combination therapies (ACTs) in the retail sector — results from focus group discussions in rural Kenya'*, PLoS ONE, 8(1), e54371. doi:10.1371/journal.pone.0054371 (open access)

Design & data: Eighteen focus group discussions across three districts — nine with retailers and nine with caregivers, deliberately kept as separate homogeneous groups. Nested inside a cluster randomised trial of subsidised paediatric ACT packs, retailer training and community awareness.

Why it is a useful example: Three teaching points in one paper: homogeneous group composition as a deliberate design decision; qualitative research used to explain why a trial succeeded and failed where it did; and group discussion surfacing market behaviour that stock data alone cannot show.

Retailers not stocking the subsidised drug cited insufficient supplies, lack of capital and staff turnover; caregivers admitted stopping the course early to save tablets for the next fever. Two supply chain failures — one upstream, one at the point of consumption — in the same eighteen discussions.

Observation and Documents: The Underused Half

Cases — triangulation, and archives as data

Observation: watching beats asking

Aseyo et al. (2018), Kisumu, Kenya

- A census of all 16 community health volunteers in one informal settlement, each studied through structured observation, a questionnaire, an in-depth interview and two focus groups.
- **Structured observation showed the number of households contacted per day was arithmetically insufficient to cover the assigned caseload — a finding produced by watching, not asking.**
- When the population is small enough, you can study all of it and skip the sampling debate.
- Human Resources for Health, 16, Article 53. doi:10.1186/s12960-018-0318-4

Documents: a coding frame across 121 years

Cockerill & Hagerman (2020), Kenya

- Historical documentary analysis of policy, legislation and governance records from 1895 to 2016, combined with expert interviews.
- **Five pre-declared attributes were coded across the whole archive, producing a defensible periodisation into five governance eras — a coding frame transparent enough for a reader to audit.**
- Documents are data, not background. Their silences are findings too.
- Ecology and Society, 25(2), Article 15. doi:10.5751/ES-11409-250215

Prolonged engagement across sources is the strongest single move available to a qualitative researcher — Okello et al. (2019, Malaria Journal, doi:10.1186/s12936-019-3061-y) spent 18 months combining observation, records and interviews at three system levels.

Ethics, Access and the Rules That Actually Apply to You

Ethics is a design activity, not a form you complete afterwards

The four recurring principles

- **Informed consent** — written or verbal, and specific: to be interviewed, recorded, photographed, or to have records accessed
- **Confidentiality and anonymity** — no detail that could identify a participant, including by combination
- **Vulnerable groups** — consent becomes more complex, especially where images are involved
- **Cross-cultural research** — in some settings it is difficult to refuse a researcher; imported ethical rules rarely transfer unmodified

Harder cases worth discussing

- Consent in observational and ethnographic fieldwork — who exactly consented?
- Consent for internet and social media data — public does not mean fair game
- Whether any deception can be appropriate — and what you lose by revealing your full interest
- Paying participants — inducement versus fair compensation
- The unintended effects of good ethical practice — participants performing moral adequacy because they are being observed

Activity — Design One Page

Activity • 12 minutes individually, then 5 minutes in threes

1	Question	Write one qualitative research question in the form “How do...” or “In what ways...”. Use the puzzle you named earlier.
2	Methodology	Name the methodology and write one sentence justifying it against the question — not against convenience.
3	Sample	Who or what, how many, and on what criteria? Name the sampling logic and give your information-power argument.
4	Method	Which data-generation method, and what will it get you that the alternatives would not?
5	Access & ethics	How will you gain access, what consent do you need, and what is the one ethical risk you must manage?
6	Weak point	State the weakest link in your own design. This is the question your examiner will ask first.

In threes: each person reads only their question and their weakest point. The other two ask one hard question each. No fixing, just interrogating.

Session 2 — What to Carry Forward

1	The question does the work	Design failures are almost always question failures. Fix the question first.
2	Name your methodology and meet its obligations	Calling a study grounded theory or ethnography commits you to specific practices. Do them or choose a different label.
3	Purposive means criteria, not convenience	Write down the criteria. If the real reason is access, say so as a limitation.
4	Argue sample size, do not assert it	Use saturation evidence and the five dimensions of information power.
5	Ethics and access are design constraints	Data Protection Act and your institution's process shape what is feasible — plan for them from the start.



Coffee / Tea Break

10:50 – 11:10



SESSION 3 • 11:10 – 11:50

Qualitative Analysis and Rigour

- What qualitative analysis actually involves
- The family of analytic approaches
- Thematic analysis: six phases, and what a theme is
- A worked coding example
- Trustworthiness: credibility, transferability, dependability, confirmability
- Reflexivity, audit trails and reporting standards

What is Qualitative Data Analysis?

Moving from material to an argument

- Qualitative data analysis is the range of processes by which we move from the data we have collected to some form of explanation, understanding or interpretation of the people and situations we are investigating.
- It examines the meaningful and symbolic content of data. From interview material you may be trying to identify:
 - someone's interpretation of their world, and why they hold that view
 - how they came to it, and what they have actually been doing
 - how they conveyed their situation to you — and how they classify themselves and others

The uncomfortable truth about the instrument

- In qualitative research, you are the instrument. The quality of the analysis is bounded by your training, your analytical skill and your willingness to be disconfirmed.
- It therefore requires both a scientific disposition (systematic, rigorous, critical) and a creative one: insightful, exploratory, willing to sit with ambiguity.

There is no software that does the thinking for you.

A Systematic Route Through the Data

Iterative, not linear — expect to revisit every step

1	Familiarise	Read and re-read. Transcribe or check transcripts yourself where possible. Note first impressions separately from the data.
2	Code and index	Attach labels to segments of text, image or interaction that mark them as instances of some idea. Index participants and sources.
3	Categorise and reduce	Group codes; abstract; collapse duplicates. This is where most of the intellectual work happens.
4	Compare	Within case and across cases. Look for patterns, and equally for differences between subgroups.
5	Build and integrate themes	Move from categories to themes with a central organising concept. Check each theme against the whole data set.
6	Refute	Actively look for data that contradicts your emerging theme. Be prepared to discard themes you like.
7	Interpret	Explain why the pattern is there. Connect to theory and to practical implications. This is the finding.

The Family of Analytic Approaches

Most involve some form of thematic work — but they are not interchangeable

Thematic analysis

- Identifying, analysing and reporting patterns across a data set.
- The most widely used approach; flexible across theoretical positions.

Content analysis

- Systematic coding against a defined frame; can report frequencies.
- Suits documents, policy and large multi-site material.

Framework analysis

- A matrix of cases by category, built from an explicit framework.
- Strong for team-based and policy-facing work.

Grounded theory coding

- Open, axial and selective coding with constant comparison.
- Aims at theory, not description.

Narrative analysis

- Analyses whole stories: structure, sequence, and how the teller positions themselves.

IPA

- Idiographic, case-by-case examination of lived experience before any cross-case claim.

Discourse analysis

- How language constructs objects, subjects and power. Does not necessarily produce themes at all.

Conversation analysis

- The fine-grained mechanics of talk in interaction: turn-taking, repair, sequence.

Do not write “thematic analysis” because it is safe. Write it because it answers your question — and then do it properly.

Thematic Analysis: The Six Phases

Braun & Clarke (2006), and their later reflexive reformulation

1	Familiarisation	Immersion in the data. Reading with analytic curiosity, making early notes — this phase cannot be delegated.
2	Generating initial codes	Systematic coding of features of the data relevant to your question, across the entire data set.
3	Searching for themes	Collating codes into candidate themes and gathering all data relevant to each.
4	Reviewing themes	Checking candidate themes against the coded extracts and then against the full data set. Themes will collapse and split.
5	Defining and naming	Writing a clear definition of each theme and what it does. If you cannot define it in a sentence, it is not a theme.
6	Producing the report	Selecting compelling extracts, relating the analysis back to the question and to the literature.

State your theoretical position explicitly. Braun & Clarke's later work stresses that themes are generated by the researcher, not discovered lying in the data — which is precisely why your process must be visible.

Braun, V. & Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3(2), 77–101. doi:10.1191/1478088706qp063oa • Braun, V. & Clarke, V. (2019) 'Reflecting on reflexive thematic analysis', *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. doi:10.1080/2159676X.2019.1628806

Two Coding Decisions You Must Declare

Both are legitimate; being unclear about which you did is not

Inductive or deductive?

Where the codes come from

- **Inductive (bottom-up):** codes are developed from the data; the research question can evolve through coding.
- **Deductive (top-down):** codes come from existing theory or a pre-set frame; you code for a specific question.
- In practice most studies are hybrid — a starting frame, refined by what the data forces you to add.
- Say which, and say when the frame changed.

Semantic or latent?

How deep the reading goes

- **Semantic:** themes are identified in the explicit, surface content — what the participant actually said.
- **Latent:** the analysis examines the assumptions, ideas and ideologies that shape what was said.
- Braun & Clarke's image: the semantic reading describes the surface of the jelly; the latent reading asks what gave it that shape.
- Latent analysis is already theorised — it is not just description.

A coding frame is not a finding. The finding is what the pattern of codes lets you argue.

A Worked Example: Coding an Extract

Illustrative teaching transcript — read it before looking at the codes

“When you move into your own home, you're alone. There is no bustle of people around the house. I miss having someone to chat to when I get home. I put the TV or some music on so there's some background noise; the silence makes me feel so alone. Sometimes I will be sat watching trash TV and thinking I should be out doing something rather than watching this rubbish. I read a lot but sometimes I am too tired and just want to veg out. But it's been good to move out of mum and dad's as it's not healthy to rely on them as they won't last forever. I became independent and made my own decisions. It's good they're still there when I need them. It's good to have some distance as when I was at home, I was arguing a lot with my dad and that was what made me decide it was time to go.”

Before the next slide: what would you code here? Write down three codes — and notice whether they are descriptions of content or interpretations of what is going on.

The Same Extract: Semantic and Latent Readings

One extract, two levels of claim

Data segment	Semantic code	Latent reading
“you're alone... no bustle of people”	Absence of company	Independence is experienced as loss as well as gain
“I put the TV or some music on... the silence makes me feel so alone”	Managing silence	Active work is required to make solitude tolerable
“I should be out doing something rather than watching this rubbish”	Self-criticism about leisure	An internalised standard of how a young adult ought to spend time
“it's not healthy to rely on them as they won't last forever”	Justifying leaving home	Independence framed as a moral obligation, not just a preference
“It's good they're still there when I need them”	Continued support available	Autonomy is claimed while dependence is quietly retained
“arguing a lot with my dad... time to go”	Conflict as trigger	The decision is narrated as inevitable rather than chosen

Candidate theme: “Independence as an achievement that must be continually justified.” Note that it is a claim about the data, not a summary of it — and it now needs testing against every other transcript.

What Counts as a Theme?

The single most common source of weak qualitative findings

Not a theme

- X A topic summary — “Communication”, “Challenges”, “Training” — a bucket that data was tipped into
- X Your interview questions rearranged as headings
- X A count — “eight participants mentioned cost”
- X A single vivid quote promoted to a finding
- X A code with a longer name

A theme

- ✓ Has a central organising concept — one idea that holds the extracts together
- ✓ Can be defined in a sentence, and that definition survives contact with the whole data set
- ✓ Says something about the data in relation to your research question
- ✓ Is evidenced across cases, and its limits are stated
- ✓ Can be argued against — and you have looked for the counter-evidence

Test each theme out loud: “My claim is that... and the evidence is... and what would count against it is...”

Rigour: Translating Validity and Reliability

The concepts still matter — the criteria have to be reworked for interpretive work

- **Validity asks whether a study can actually answer the question it claims to answer.**
- **Reliability asks about consistency** — your bathroom scales may be perfectly consistent and still wrong, which is why reliability alone guarantees nothing.
- **Triangulation compares evidence from more than one method, source, researcher or theory, so a claim does not rest on a single reading.**
- In interpretive research these ideas cannot be applied unchanged: there is no expectation that a second researcher would produce an identical account, because the researcher is part of the instrument. Lincoln and Guba's four criteria reframe the question as trustworthiness, and attach practical techniques to each:

Criterion	The question it answers	How you demonstrate it
Credibility	Is the account a defensible reading of the participants' reality?	Prolonged engagement, triangulation, peer debriefing, negative case analysis, member checking, verbatim evidence
Transferability	Could a reader judge whether this applies to their context?	Thick description of setting, participants and boundaries; explicit statement of scope
Dependability	Is the process traceable and documented?	Audit trail of decisions, coding frame versions, memos, dated analytic notes
Confirmability	Are the findings shaped by the data rather than by the researcher's preferences?	Reflexive account of positionality, disconfirming evidence reported, raw-data-to-claim chain shown

Lincoln, Y.S. & Guba, E.G. (1985) *Naturalistic Inquiry*. Newbury Park, CA: Sage.

Eight Markers of Excellent Qualitative Research

A broader, cross-tradition quality vocabulary — Tracy (2010)

Worthy topic

- Relevant, timely, significant — the study is worth doing

Rich rigour

- Sufficient, appropriate and abundant data, theory, samples and time in the field

Sincerity

- Self-reflexivity about values and biases; transparency about method and challenges

Credibility

- Thick description, triangulation, multivocality, member reflections

Resonance

- The work affects readers through evocative writing and transferable findings

Significant contribution

- Theoretical, practical, moral, methodological or heuristic contribution

Ethics

- Procedural, situational, relational and exiting ethics — not just the approval form

Meaningful coherence

- The study achieves what it claims, and method, literature and findings hang together

Tracy, S.J. (2010) 'Qualitative quality: eight "big-tent" criteria for excellent qualitative research', *Qualitative Inquiry*, 16(10), 837–851. doi:10.1177/1077800410383121

Six Techniques You Can Actually Report

Each one produces documentary evidence a reviewer can inspect

1	Triangulation	Across methods, sources, researchers or theories. Report where accounts converged and, more usefully, where they did not.
2	Audit trail	Dated record of analytic decisions, coding frame versions and memos. This is what makes dependability demonstrable rather than asserted.
3	Reflexivity	A written positionality account: who you are in relation to the field, and how that shaped access, questions and interpretation.
4	Peer debriefing	A colleague independently reads extracts and challenges your coding. Report the disagreements and how they were resolved.
5	Negative case analysis	Deliberately search for data that contradicts each theme, and revise the theme. Report what you found.
6	Member checking	Return analysis to participants — but treat it as generating further data, not as validation.

On member checking: Birt et al. (2016) caution that returning findings can be “merely a nod to validation”. Their alternative, synthesised member checking, returns the synthesised analysis rather than raw transcripts — and treats participants’ responses as new data.

Birt, L., Scott, S., Cavers, D., Campbell, C. & Walter, F. (2016) 'Member checking: a tool to enhance trustworthiness or merely a nod to validation?', *Qualitative Health Research*, 26(13), 1802–1811.
doi:10.1177/1049732316654870

Reflexivity in Practice

Not a confession — a methodological instrument

Before fieldwork

- What do I already believe I will find, and where did that belief come from?
- What is my relationship to this setting — insider, outsider, or something more awkward?
- Which findings would be inconvenient for me, my funder or my institution?

During fieldwork

- How is my presence changing what I am seeing?
- Who is talking to me easily, and who is not — and why?
- What am I not asking about because it feels uncomfortable?

During analysis

- Whose account am I privileging, and on what grounds?
- Which extracts am I drawn to, and is that because they are typical or because they are quotable?
- What is the strongest alternative reading of this data, and why am I rejecting it?

Keep this in a dated research journal from day one. It becomes your audit trail, your reflexivity statement and, eventually, several paragraphs of your methods chapter.

Reporting Standards and Appraisal Checklists

Write to the checklist your reviewers will use

COREQ

32 items — interviews and focus groups

- Three domains: the research team and reflexivity; study design; analysis and findings.
- Widely required by health and social science journals.
- Tong, A., Sainsbury, P. & Craig, J. (2007), International Journal for Quality in Health Care, 19(6), 349–357.
doi:10.1093/intqhc/mzm042

SRQR

21 items — qualitative research generally

- Broader than COREQ; not restricted to interview and focus group designs.
- Useful when your design is ethnographic, documentary or mixed.
- O'Brien, B.C. et al. (2014), Academic Medicine, 89(9), 1245–1251.
doi:10.1097/ACM.0000000000000388

CASP checklists

For appraising what you read

- Free downloadable critical appraisal checklists for several study types, including qualitative studies.
- Use them on the papers in your literature review — then on your own draft.
- casp-uk.net/casp-tools-checklists/

Practical move: download COREQ or SRQR now and keep it open while you write your methods chapter. It converts “be rigorous” into a list of things to state.

A Published Example: Rigour Made Visible

Case — what a defensible process looks like on the page

PUBLISHED EXAMPLE

Kimani-Murage, E.W., Osogo, D., Nyamasege, C.K., Igonya, E.K., Ngira, D.O. & Harrington, J. (2022) *'COVID-19 and human right to food: lived experiences of the urban poor in Kenya with the impacts of government's response measures, a participatory qualitative study'*, BMC Public Health, 22, Article 1399. doi:10.1186/s12889-022-13638-3 (open access)

Design & data: Korogocho and Viwandani informal settlements, Nairobi. Eighty participants across eight focus groups (deliberately homogeneous: women, men, youth, traders), eight in-depth interviews, twelve key informant interviews, two photovoice sessions and three digital storytelling sessions. Thematic analysis organised by an a-priori right-to-food framework, with dual coding and documented code agreement across the team.

Why it is a useful example: Five data-generation formats, an explicitly named analytic framework declared in advance, and dual coding with reconciliation — three separate trustworthiness moves, each one stated plainly enough that a reader could challenge it. That is what dependability and confirmability look like in a published methods section.

Notice also that a named philosophical tradition is operationalised rather than merely name-dropped.

Activity — Code and Defend

activity in pairs

1	Code independently	Take the extract from earlier in this session. Each person codes it alone, in silence, for three minutes.
2	Compare	Exchange codes. Where did you differ? Disagreement is the interesting part, not a failure.
3	Escalate one code	Choose one semantic code and push it to a latent reading. What are you now claiming beyond what was said?
4	Name a candidate theme	Write one theme with a central organising concept, in one sentence.
5	Attack it	Your partner's job: state what evidence would defeat that theme. Yours: say how you would look for it.

This exercise is peer debriefing and negative case analysis in miniature — the two techniques most often promised in methods chapters and least often actually done.

Session 3 — What to Carry Forward

1	Coding is not analysis (but is part of it)	Codes organise data. The finding is the argument you can defend about why the pattern exists.
2	A theme needs a central organising concept	If it is a topic bucket or a rearranged interview question, it is not yet a theme.
3	Rigour is documented, not claimed	Credibility, transferability, dependability and confirmability each have techniques that leave a paper trail.
4	Look for what would defeat you	Negative case analysis and reflexivity are what separate interpretation from confirmation.
5	Write to a standard	COREQ or SRQR turns “rigorous” into a checkable list before a reviewer does it for you.

SESSION 4 • 12:00 – 13:00

Tools for Qualitative Research and the Role of AI

- What software can and cannot do for a qualitative researcher
- The 2026 landscape: NVivo, ATLAS.ti, MAXQDA, Dedoose, Quirkos, Taguette, QualCoder
- Choosing a tool on methodological grounds
- What the peer-reviewed evidence says about AI-assisted analysis
- Bias, reproducibility and fabricated quotations
- Law, disclosure and a working protocol for responsible use

The tool serves the method. Never the reverse.

This session is not software training. It is about making tool choices you can defend in a methods chapter — and knowing where AI breaks rigour.

What Qualitative Software Actually Does

Set expectations correctly and the tool becomes genuinely useful

What it does well

- ✓ Centralises heterogeneous material — transcripts, notes, documents, images, audio, video — in one retrievable place
- ✓ Makes coding systematic and revisable: rename, merge, split, and see every extract at a code in one view
- ✓ Supports querying: matrices by case and code, text search, co-occurrence, comparison across subgroups
- ✓ Keeps an audit trail of analytic decisions — direct evidence for dependability
- ✓ Enables genuine team coding with visible disagreement rather than silent divergence
- ✓ Extends the life and reusability of a dataset beyond one project

What it does not do

- X It does not decide what is interesting, or what your data means
- X It does not make a weak research question workable
- X It does not turn a topic bucket into a theme
- X Autocoding by pattern or heading organises text; it does not interpret it
- X It does not confer rigour — an unjustified design coded neatly is still unjustified
- X It does not remove your obligation to know the analytic approach you are claiming to use

You can produce excellent qualitative research with paper, highlighters and a spreadsheet. Software changes the scale and the auditability, not the thinking.

The 2026 Landscape

Four categories, following the CAQDAS Networking Project's own classification

Traditional CAQDAS

Full-function analysis environments

- NVivo, ATLAS.ti, MAXQDA, Dedoose, Quirkos, HyperRESEARCH, f4Analyse, webQDA
- Code, retrieve, query, memo, visualise, work in teams
- Most now embed generative-AI features alongside conventional tools

Method-oriented tools

Built around one approach

- Leximancer, QDA Miner, WordSmith Tools, DiscoverText, Transana, MiMeG
- Content analysis, corpus linguistics, text mining, video and multimodal analysis
- Powerful within their method; awkward outside it

GenAI-driven tools

AI-first, newer, volatile

- A fast-moving category; products rebrand and change terms frequently
- Often promise thematic analysis “automatically”
- Treat vendor accuracy claims as claims to be tested, not facts

Free and open source

Realistic where budgets are tight

- Taguette — free, open source (BSD), local install or hosted; import, highlight, tag, export
- QualCoder — free, open source (LGPLv3); codes text, images, audio and video
- No licence cost, no vendor lock-in, and your data need never leave your machine

Taxonomy and independent software reviews: CAQDAS Networking Project, University of Surrey — surrey.ac.uk/computer-assisted-qualitative-data-analysis. Note their reviews lag current vendor releases; read them for analytic capability rather than version detail. Verified 6 August 2026.

Seven Tools, Compared on What Matters to You

Platform and AI posture — verified from vendor pages, August 2026

Tool	AI posture	Notes for a student
NVivo (Lumivero)	AI Assistant sold as an add-on; summaries and sub-code suggestions you accept, refine or dismiss	The most commonly taught package; vendor states AI is designed to support, not replace, the researcher
ATLAS.ti	AI coding using a commercial LLM; from v26 an MCP server lets you connect your own approved AI client locally	The bring-your-own-model route is the most privacy-controllable of the commercial options
MAXQDA (VERBI)	AI Assist with a free limited tier; optional processing restricted to EU servers	The EU-only processing option matters if your data cannot lawfully leave a jurisdiction
Dedoose	No AI features advertised	Cheapest credible commercial route; browser-based, so no install
Quirkos	Explicitly no AI analysis; states it does not use your data to train AI	Runs on Windows, Mac, Linux and in a browser, including tablets
Taguette	None	Import, highlight, tag, export. Local install or hosted. Ideal for a first coded project
QualCoder	Optional AI using your own API key, or a free academic model service; local analysis supported	Codes text, images, audio and video. The strongest genuinely free AI-capable option

Compiled from lumivero.com, atlasti.com, maxqda.com, dedoose.com, quirkos.com, taguette.org and the QualCoder repository, verified 6 August 2026. Prices and features change frequently — check before you commit.

Choosing a Tool: Six Questions

Answer these before you look at a feature list

1	What data types?	Text only, or images, audio, video, social media, survey open-ends? This eliminates most options immediately.
2	Alone or in a team?	Team coding, merged projects and visible disagreement need specific features — and usually a paid licence.
3	What is your real budget?	Including after your course ends. A project you cannot open in two years is a problem.
4	Where may your data legally live?	Cloud processing may constitute a cross-border transfer of personal data. Local or in-region processing may be required.
5	Can you get your data out?	Check the export formats before you import anything. Lock-in is the most common regret.
6	What will you have to learn?	Time spent learning software is time not spent analysing. Match the tool to the size of the project.

For a masters dissertation with 15 interviews, Taguette or Dedoose is often the rational choice. Scale the tool to the project, not to the brochure.

Where AI Plausibly Sits in a Qualitative Workflow

Before, alongside and after the analysis — never instead of it

Before analysis

- Cleaning and formatting anonymised transcripts
- Drafting an interview or discussion guide you then critique
- Developing sensitising concepts from the literature
- Generating a first-draft codebook you expect to substantially rewrite

Alongside analysis

- Stress-testing an interpretation: “what is the strongest alternative reading?”
- Devil's advocate on an emerging theme
- Drafting analytical memos you then edit and own
- Suggesting codes you may have missed — as prompts, not verdicts

After analysis

- Structuring the findings narrative you have already reasoned out
- Testing whether your theme definitions are actually clear to a reader
- Surfacing links to theory for you to verify against real sources
- Generating counterarguments for the discussion section

Note what is absent from all three columns: “decide what the data means”. Every use above keeps the interpretive judgement — and the accountability — with you.

The Evidence, 1: There Is No “Accuracy” Figure

The most useful number in this session is a range, not a value

“Agreement between large language model and human coders across the published empirical literature ranges from 36% to 99%.”

Kempny, C. et al. (2026) 'The use and methodological reporting of large language models in qualitative research: a scoping review', BMC Medical Research Methodology, Article 137. doi:10.1186/s12874-026-02913-1 — 5,049 records screened, 75 empirical studies included.

What else that review found

- GPT-family models were used in 93% of studies; open-source models were used notably less often.
- Most common applications: coding assistance and theme identification; most common analytic frames: thematic and content analysis.
- Reporting is poor. 75% of the studies reported no model parameter settings at all; 45% did not specify the deployment configuration.
- Almost all included some form of human verification — which tells you what the field itself thinks is required.

Why this matters for your own work

- A vendor claim such as “reduce analysis time by more than 90%” or “90%+ accuracy” is a marketing statement, not a finding. Ask: accuracy against what standard, on whose data, measured how?

The Evidence, 2: Fast at the Surface, Weak at Depth

The trade-off is consistent across independent studies

Two hours instead of twenty-three

Morgan (2023)

- Compared ChatGPT against the author's own reflexive thematic analysis and iterative thematic inquiry on two datasets.
- **Roughly 2 hours with the model versus about 23 hours of manual analysis.**
- **But: more successful at reproducing concrete, descriptive themes; less successful at locating subtle, interpretive ones.**
- Example: it produced “maintaining a balance between school and social life” where the human analysis reached the overarching “navigating the transition to a new role”.
- International Journal of Qualitative Methods, 22. doi:10.1177/16094069231211248 (open access)

More codes, shallower codes

Hitch (2023)

- A worked comparison positioning AI as a “virtual colleague” augmenting reflexive thematic analysis.
- **On the same material the model produced 10 codes against the researcher's 5, and 16 against 7.**
- **Its codes were markedly more literal; the researcher's engaged latent meaning.**
- The comparison was still useful — it provoked deeper reflexivity about why she had coded as she did.
- Qualitative Health Research. doi:10.1177/10497323231217392 (open access)

Read together: AI is fastest and most reliable precisely where the analytic value is lowest.

The Evidence, 3: The Quotations Problem

The failure mode that will damage your thesis most — evidenced on Kenyan data

PUBLISHED EXAMPLE

Mehta, S.D., Paul, S., Awiti, E., Young, S., Zulaika, G., Otieno, F.O., Phillips-Howard, P.A., Mason, L., Bhaumik, R. (2025) 'Evaluation of large language models within GenAI in qualitative research', *Scientific Reports*, 15, Article 34993. doi:10.1038/s41598-025-18969-w (open access)

Design & data: GPT-4o compared against human thematic analysis of focus group transcripts on COVID-19 and sexual and reproductive health among adolescent girls and young women in rural Kenya.

Why it is a useful example: The theme labels looked right — 13 themes from the female transcripts, matching the human analysis, with no disagreement on sub-themes. The evidence underneath them did not hold. Only 64–87% of quotes for female transcripts and 33–79% for male transcripts were rated consistent with the themes they were offered for; strongly supportive quotes fell to 36–67% and 18–55% respectively. Quotes were also modified — from a single word change through to truncation and recombination that altered meaning.

The authors' conclusion: the model “was unable to provide a thematic analysis that is indistinguishable from a human analysis”. Note also that performance was systematically worse on male transcripts — differential accuracy across subgroups, on African data.

Plausible themes with unverifiable evidence is the worst possible outcome: it looks finished.

The Evidence, 3 continued: Two More Independent Findings

Fabricated evidence, and a model that will confirm a framework that does not fit

It cited a trip that was never mentioned

Hamilton et al. (2023)

- 71 guaranteed-income interviews; 1,125 significant statements analysed by humans using Colaizzi's method and by ChatGPT.
- Humans produced 11 themes, the model 10, with overlap on about half the human themes and 80% of the model's.
- **The model cited “a trip to Niagara Falls” as evidence — the phrase appeared exactly once across all interviews.**
- It missed family motivation, faith, privacy, asset development and personal development, and never named inflation.
- IJQM, 22. doi:10.1177/16094069231201504 (open access)

It validated a framework that did not fit

Wachinger et al. (2024)

- One 58-minute interview, GPT-3.5 against an experienced researcher using reflexive thematic analysis.
- **In two cases the quotes the model included did not exist in the transcript; others were edited without any flag.**
- **Given a deliberately unsuitable theoretical framework, it argued convincingly that the themes fitted — and proceeded without ever flagging the poor fit.**
- Codes and quotes had been retrofitted to themes, and it could not explain its own analytic process.
- Qualitative Health Research. doi:10.1177/10497323241244669 (open access)

Practical rule with no exceptions: every quotation that appears in your work is verified against the original transcript by you, by search, one at a time.

The Evidence, 4: The Errors Are Not Random

This is the finding that should concern researchers working in the Global South most

PUBLISHED EXAMPLE

Ashwin, J., Chhabra, A. & Rao, V. (2025) 'Using large language models for qualitative analysis can introduce serious bias', Sociological Methods & Research. doi:10.1177/00491241251338246

Design & data: 789 interview transcripts — 9,964 question-and-answer pairs — from Rohingya refugees and Bangladeshi host communities in Cox's Bazaar. Four large language models compared against expert human annotation, and against a small purpose-trained supervised model.

Why it is a useful example: Two findings matter. First, absolute performance was modest: mean F1 of 0.414 for the best general model, against 0.542 for the small supervised model trained on the actual data. Second, and more seriously, for 10 of 19 codes the model's errors correlated significantly with respondent characteristics — refugee status, gender, education and income. Sparse codes were over-predicted: humans found secular aspirations 1.2 times more frequent than religious ones, while the model reported three times.

Errors that correlate with participant characteristics do not average out. They systematically distort exactly the comparisons between groups that most social research is trying to make — and only expert human annotation made the bias detectable at all.

The Evidence, 5: The Same Prompt, A Different Answer

Reproducibility and the black box

Seven themes in February, five in October

Kabir et al. (2025)

- Manual reflexive thematic analysis produced 9 themes per participant; an AI tool produced 7.
- **Repeating the identical analysis on the identical transcripts eight months later produced only 5 themes per transcript.**
- It missed “Supervision” entirely and initially missed “Role Satisfaction and Professional Growth”.
- IJQM, 24. doi:10.1177/16094069251336810

Four rigour problems, named

Reeping, Hampton & Özkan (2025)

- **Ethical validation — uploading participant data to commercial models without consent can breach your ethics approval.**
- **Theoretical validation — training data reflects dominant Global North narratives, risking automated misrepresentation of other contexts.**
- **Procedural validation — black-box reasoning cannot be inspected, so your analytic chain is not auditable.**
- **Process reliability — “there is no guarantee that even the same prompt will produce the same output every time”, compounded by silent model updates.**
- Studies in Engineering Education, 6(2), 1–23. doi:10.21061/see.174 (open access)

black-box reasoning defeats dependability; unstable output defeats confirmability; undisclosed data transfer defeats research ethics.

So Where Does That Leave Us?

A defensible division of labour, given the evidence

Defensible uses

- ✓ Anonymised transcript tidying and formatting
- ✓ Drafting instruments and prompts you then critique and rewrite
- ✓ Generating a candidate codebook you treat as a hypothesis, not a result
- ✓ Challenging your interpretation: alternative readings, counterarguments, gaps
- ✓ Simulating a hostile reviewer on your own draft
- ✓ Explaining unfamiliar methodological terms — then verified against a real source
- ✓ Language and clarity editing of your own prose

Indefensible uses

- X Pasting raw, identifiable transcripts into a public tool
- X Accepting AI-generated quotations without verifying each one against the transcript
- X Reporting AI-generated themes as your findings without independent analysis
- X Trusting AI-generated citations — verify every one in Scopus, Web of Science or Google Scholar
- X Using AI to reach a conclusion you then present as human reasoning
- X Letting a model tell you your framework fits — it will agree with you
- X Using it at all without disclosing it where policy requires

The pattern: AI is acceptable where it generates something you will critically judge, and unacceptable where it replaces the judgement itself.

Publication ethics

COPE and ICMJE

- COPE (v1, August 2024): AI tools cannot be listed as an author, because they cannot take responsibility for the work.
- **COPE requires disclosure where AI was used in “the collection and analysis of data” — so AI-assisted coding belongs in your Methods section, not a footnote.**
- ICMJE: journals should require disclosure of AI-assisted technologies at submission; authors remain responsible for everything, including AI-produced parts.
- Authors are fully liable for any breach of publication ethics arising from AI use.

What Publishers Require of You

Five major publishers, five different rules — check the one you are submitting to

Publisher	What they require of authors
Elsevier	A separate AI declaration statement in the manuscript at submission; basic grammar checking exempt; no AI generation or alteration of images representing primary data
Springer Nature	Disclosure required; human accountability cannot be transferred to AI; no photorealistic generated images; no presenting AI-generated analyses or conclusions as human-derived
Emerald	The strictest: generating any new text with AI is not permitted at all. Copy-editing your own writing is. Disclose in both Methods and Acknowledgements, with tool name and version
Taylor & Francis	Disclose the full tool name with version number, how it was used and why; for books, disclose to your editor in advance and state it in the preface
Sage	Assistive use (grammar, structure) needs no disclosure; generative use affecting methodology, analysis, results or conclusions must be disclosed in Methods or Acknowledgements. Conducting interviews with AI in place of participants is prohibited

publicationethics.org/cope-position-statements/ai-author • icmje.org • elsevier.com • nature.com/nature-portfolio/editorial-policies/ai • emeraldgroupublishing.com • taylorandfrancis.com/our-policies/ai-policy • sagepub.com. Verified 6 August 2026. Policies change — check the current version for your target journal.

A Working Protocol You Can Adopt Today

Eight rules that keep AI use inside the boundaries of rigour, ethics and law

1	Anonymise first	Remove names, organisations, places and any identifying combination before anything touches a model. No exceptions for “quick” checks.
2	Check where processing happens	Prefer local processing, in-region servers or institutionally approved tools. Cloud processing of participant data may be a cross-border transfer.
3	Never outsource the judgement	Use AI to generate candidates and challenges. You decide what the data means, every time.
4	Verify every quotation	Search the original transcript for each quoted string. Studies have documented quotes that were altered, and quotes that never existed.
5	Verify every citation	Check each reference in Scopus, Web of Science or Google Scholar. Plausible fabricated references are the most common failure.
6	Log what you did	Record the tool, model, version, date, the prompts used and what you accepted or rejected. This is your audit trail.
7	Disclose in the Methods	COPE covers data collection and analysis. Name the tool and version and describe the use. Then check your target journal's specific wording.
8	Keep a human-only comparison	On a subset, analyse without AI and compare. It is the only way to know what the tool is doing to your analysis.

If You Do Use It: Prompt for Thinking, Not for Output

Structure the request, then use the response as an argument to interrogate

Structure a research prompt

- **Role** — “You are an experienced qualitative methodologist...”
- **Context** — the study, the design, the stage you are at
- **Task** — precisely what you want, stated as a single job
- **Format** — length, structure, headings
- **Constraints** — “Do not invent references. Do not quote material I have not given you.”
- Then iterate. Treat it as a dialogue, not a single query.

Prompts that improve your thinking

- **Devil's advocate** — “What are the strongest counterarguments to this theme?”
- **Gap finder** — “Given these five themes, what perspective am I likely missing?”
- **Reviewer simulation** — “Act as a critical reviewer for this journal and attack this abstract's contribution.”
- **Socratic** — “Ask me five questions about my design that would challenge my framework.”
- **Alternative reading** — “Argue that my data supports the opposite interpretation.”

Notice that every prompt on the right asks the model to disagree with you. That is the use with the best evidence behind it — and the one least likely to damage your rigour.

Activity — Audit an AI Output

1	Prepare	Take a short piece of anonymised text you are comfortable using.
2	Prompt	Ask a model for three themes with supporting quotations, using Role + Context + Task + Format + Constraints.
3	Audit the quotations	Search the original text for every quoted string. Mark each: exact, altered, or absent. Record the counts.
4	Audit the depth	Are the themes semantic or latent? Are they topic buckets, or do they have a central organising concept?
5	Re-prompt as a critic	Now ask it to argue against its own themes. Which response was more useful to you — and why?
6	Report back	One sentence to the room: the single thing you would and would not trust it with.

Step 3 is the point of the exercise. Most people are surprised — in both directions.

Session 4 — What to Carry Forward

1	Tools follow method	Choose software from your data types, your team, your budget and your legal constraints — then justify it in your methods chapter.
2	There is no accuracy figure	Published agreement between AI and human coders spans 36% to 99%. Treat any single vendor number as marketing.
3	Verify the evidence, not just the themes	AI theme labels can look right while the supporting quotations are altered or fabricated. Check every one.
4	Errors correlate with people	AI coding errors have been shown to track respondent characteristics — which distorts exactly the group comparisons research depends on.
5	Disclose, log, and stay legal	Keep a prompt and decision log.

The Whole Seminar in Six Sentences

1	Start with a real puzzle	Qualitative research earns its place when the question is about meaning, process or context — not when the numbers are inconvenient.
2	Make your assumptions explicit	Ontology, epistemology and axiology are already shaping your design. Declaring them is what makes the design defensible.
3	Let the question choose the method	Name your methodology, meet its obligations, and justify your sample with criteria and information power.
4	Coding is not the finding	The finding is the argument you can defend about why a pattern exists — with a theme that has a central organising concept.
5	Rigour is a paper trail	Credibility, transferability, dependability and confirmability are demonstrated through documented techniques, not asserted in a sentence.
6	Tools and AI serve the thinking	They change scale, speed and auditability. They do not do the interpretation, and they do not carry the accountability. You do.

Key Sources — Methodology

- Birt, L., Scott, S., Cavers, D., Campbell, C. & Walter, F. (2016) 'Member checking: a tool to enhance trustworthiness or merely a nod to validation?', *Qualitative Health Research*, 26(13), 1802–1811. doi:10.1177/1049732316654870
- Braun, V. & Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3(2), 77–101. doi:10.1191/1478088706qp063oa
- Braun, V. & Clarke, V. (2019) 'Reflecting on reflexive thematic analysis', *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. doi:10.1080/2159676X.2019.1628806
- Eisenhardt, K.M. (1989) 'Building theories from case study research', *Academy of Management Review*, 14(4), 532–550. doi:10.5465/amr.1989.4308385
- Glaser, B.G. & Strauss, A.L. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Guest, G., Bunce, A. & Johnson, L. (2006) 'How many interviews are enough? An experiment with data saturation and variability', *Field Methods*, 18(1), 59–82. doi:10.1177/1525822X05279903
- Lincoln, Y.S. & Guba, E.G. (1985) *Naturalistic Inquiry*. Newbury Park, CA: 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. doi:10.1177/1049732315617444
- O'Brien, B.C., Harris, I.B., Beckman, T.J., Reed, D.A. & Cook, D.A. (2014) 'Standards for reporting qualitative research (SRQR)', *Academic Medicine*, 89(9), 1245–1251. doi:10.1097/ACM.0000000000000388
- Saunders, M.N.K., Lewis, P. & Thornhill, A. (2023) *Research Methods for Business Students*, 9th edn. Harlow: Pearson.
- Silverman, D. (2017) *Doing Qualitative Research*, 5th edn. London: Sage.
- Tong, A., Sainsbury, P. & Craig, J. (2007) 'Consolidated criteria for reporting qualitative research (COREQ)', *International Journal for Quality in Health Care*, 19(6), 349–357. doi:10.1093/intqhc/mzm042
- Tracy, S.J. (2010) 'Qualitative quality: eight “big-tent” criteria for excellent qualitative research', *Qualitative Inquiry*, 16(10), 837–851. doi:10.1177/1077800410383121
- Yin, R.K. (2018) *Case Study Research and Applications: Design and Methods*, 6th edn. Thousand Oaks, CA: Sage.

Key Sources — Example Studies and AI Evidence

Published example studies (all open access)

- Aseyo, R.E. et al. (2018) 'Realities and experiences of community health volunteers...Kisumu, Kenya', Human Resources for Health, 16, 53. doi:10.1186/s12960-018-0318-4
- Botes, A., Niemann, W. & Kotzé, T. (2017) 'Buyer–supplier collaboration and supply chain resilience', South African Journal of Industrial Engineering, 28(4), 183–199. doi:10.7166/28-4-1736
- Cockerill, K.A. & Hagerman, S.M. (2020) 'Historical insights...community-based conservation in Kenya', Ecology and Society, 25(2), 15. doi:10.5751/ES-11409-250215
- Hennegan, J. et al. (2020) "'I do what a woman should do"...Mukono District, Uganda', BMJ Global Health, 5(11), e003433. doi:10.1136/bmjgh-2020-003433
- Jepkosgei, J., English, M., Adam, M.B. & Nzinga, J. (2022) 'Understanding intra- and interprofessional team...Kenyan hospitals', BMC Health Services Research, 22, 636. doi:10.1186/s12913-022-08039-6
- Kedenge, S.V. et al. (2013) 'Understanding the impact of subsidizing ACTs in the retail sector...rural Kenya', PLoS ONE, 8(1), e54371. doi:10.1371/journal.pone.0054371
- Kimani-Murage, E.W. et al. (2022) 'COVID-19 and human right to food...Kenya', BMC Public Health, 22, 1399. doi:10.1186/s12889-022-13638-3
- Okello, G. et al. (2019) 'Producing routine malaria data...frontline health facilities in Kenya', Malaria Journal, 18, 420. doi:10.1186/s12936-019-3061-y

Peer-reviewed evidence on AI in qualitative analysis

- Ashwin, J., Chhabra, A. & Rao, V. (2025) 'Using large language models for qualitative analysis can introduce serious bias', Sociological Methods & Research. doi:10.1177/00491241251338246
- Hamilton, L., Elliott, D., Quick, A., Smith, S. & Choplin, V. (2023) 'Exploring the use of AI in qualitative analysis', IJQM, 22. doi:10.1177/16094069231201504
- Hitch, D. (2023) 'Artificial intelligence augmented qualitative analysis: the way of the future?', Qualitative Health Research. doi:10.1177/10497323231217392
- Kabir, S.M.A., Ali, F., Ahmed, R.L. & Sulaiman-Hill, R. (2025) 'Exploring the use of AI in qualitative data analysis: comparing manual processing with Avidnote for theme generation', IJQM, 24. doi:10.1177/16094069251336810
- Kempny, C., Frings, J., Rust, P., Meister, S. & Fehring, L. (2026) 'The use and methodological reporting of large language models in qualitative research: a scoping review', BMC Medical Research Methodology, Article 137. doi:10.1186/s12874-026-02913-1
- Mehta, S.D., Paul, S., Awiti, E., Young, S., Zulaika, G., Otieno, F.O., Phillips-Howard, P.A., Mason, L. & Bhaumik, R. (2025) 'Evaluation of large language models within GenAI in qualitative research', Scientific Reports, 15, Article 34993. doi:10.1038/s41598-025-18969-w
- Morgan, D.L. (2023) 'Exploring the use of AI for qualitative data analysis: the case of ChatGPT', IJQM, 22. doi:10.1177/16094069231211248
- Reeping, D., Hampton, C. & Özkan, D. (2025) 'Interrogating the use of large language models in qualitative research...', Studies in Engineering Education, 6(2), 1–23. doi:10.21061/see.174
- Wachinger, J. et al. (2024) 'Prompts, pearls, imperfections: comparing ChatGPT and a human researcher in qualitative data analysis', Qualitative Health Research. doi:10.1177/10497323241244669

Thank you

Questions, Comments, Thoughts

Dr. Simon Peter Nadeem

simon.nadeem@kuehne-foundation.org

LEARN Logistics by Kuehne Foundation