FROM THE RESEARCH BENCH

A case study is an evidence design, not a story.

A practical case study research design: set the boundary, test rivals, plan saturation, and state exactly what the evidence can transfer.

Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher

PUBLISHED
23 August 2026

READING TIME
11 min

WORDS
2,316

https://isoglu.com/insights/journal/case-study-is-an-evidence-design/

CONTENTS

MANAGEMENT SUMMARY

A case study becomes useful evidence when it declares what the case is, what it is not, which unit and time window are being studied, why the case was selected, and how rival explanations will be tested. This practical case study research design brings Eisenhardt's iterative process together with the Gioia data-structure logic and four conditional approaches to sample adequacy. Guest's codebook result, Hennink's distinction between code and meaning saturation, Hagaman's cross-site metathemes, and Malterud's information power are planning anchors, not universal interview counts. The piece ends with a usable boundary checklist and a plain limit: a purposive case can explain a mechanism or condition, but it does not become a population-level causal estimate merely because it has a named company, several interviews, or a long reference list.

Keywords: Case study research · Research design · Qualitative rigor · Evidence quality · Mixed methods

A case study can look rigorous from a distance. It has a company name, a chronology, a few interviews, and a reference list long enough to quiet the room. None of those features answers the first question a reader should ask: what exactly is the case, and what would count against the explanation?

That is the distinction this piece makes. A case study is not evidence because it is vivid. It is evidence when the design makes its boundary, unit, time window, selection logic, sources, analysis, rivals, and transfer conditions inspectable. This is the same discipline behind the evidence threshold: the document matters less than what the claim had to survive before it reached the page.

Why is an organization name merely a research setting rather than a case boundary?

Suppose a growth team writes a case about a new-market launch. The company may be the setting. The case might instead be the launch decision, the first three country entries, the buyer-seller relationship, or the sequence by which a channel conflict changed the plan. Each choice creates a different unit, evidence set, and transfer claim.

The practical test is simple: write the case boundary before you collect the most interesting stories. State what enters the case, what stays outside it, when the observation starts and ends, and which outcome or mechanism the case is meant to illuminate. The matrix below turns that test into a working object. Its left column is what a narrative often leaves implicit. Its right column is what an evidence design has to show.

Table 1 From narrated case to evidence design

Design field	Story version	Evidence-design version
Question	What happened at this company?	Which decision, mechanism, or condition is being examined?
Case boundary	The company, brand, or project name	Explicit inclusion and exclusion rules, with a start and end point
Unit	The organisation treated as one actor	The decision, relationship, event, team, site, or organisation that answers the question
Selection	The most visible success or the easiest access	A stated theoretical reason: replication, extension, conceptual category, or contrasting type
Evidence	Interviews that make the chronology readable	Interviews, observation, documents, system traces, or measures selected for the question, with provenance
Analysis	A smooth chronology and a list of themes	Within-case reconstruction, cross-case or cross-source comparison, and searches for reversals
Rival explanation	A footnote that names one alternative	A log of what the rival predicts, which observations support it, and what remains unresolved
Transfer	This is what works	These conditions may travel; these conditions are local; this magnitude is not identified

Author's operationalisation of Eisenhardt (1989) and Gioia, Corley & Hamilton (2013). The right column is a design aid, not a published measurement instrument.

The hardest row is usually selection. A case chosen because it succeeded is not automatically invalid. It becomes a problem when the selection is allowed to masquerade as a typical outcome. Theoretical sampling gives the choice a reason. You may select a case that should reproduce an emerging pattern, extend it into a different condition, fill a conceptual category, or expose a contrasting type. That logic does not create statistical representativeness. It tells the reader what the case was chosen to teach.

The same applies to the evidence column. A customer interview can explain an interpretation. A CRM trace can show sequence. A contract can settle what was promised. A lost deal, a rejected proposal, or a contradictory internal memo can carry more weight against a polished explanation than another supportive interview. Triangulation is not the accumulation of agreeable voices. It is the deliberate assembly of sources that can disagree.

What does Eisenhardt's case study methodology actually prescribe for theory building?

Eisenhardt's 1989 article is useful because it treats case-based theory building as a movement between question and evidence. The process starts with a broad research question and a provisional conceptual orientation. It then uses theoretical case selection, multiple data methods or investigators where useful, overlapping collection and analysis, within-case familiarity, cross-case pattern searches, comparison with literature, and a closure judgement when incremental theoretical learning becomes small.

That sequence changes how a commercial case is written. You do not wait until the end to discover what the evidence means. You record the first explanation, collect the evidence it predicts, search for a case or source that could reverse it, and revise the boundary if the question has changed. The revision is not a defect if it is recorded. An unrecorded revision is how a case quietly becomes a story about what the team already believes.

Gioia, Corley, and Hamilton add a useful visibility layer for inductive work. First-order codes keep the informant's terms close to the material. Second-order concepts move toward the researcher's theoretical language. Aggregate dimensions show the higher-level structure. A data structure lets a reader see how the conclusion was built rather than asking them to trust the author's fluency.

But Gioia et al. explicitly position the approach as a flexible orientation, not a cookbook. That warning matters in a commercial setting, where a framework can become a badge. If your interviews do not support a three-level structure, do not force them into one. If system records contradict the interview account, do not call the contradiction noise until you have explained why it is outside the case.

Why does qualitative evidence saturation resist arbitrary interview quota targets?

The question how many interviews are enough is usually asked too late. Before naming a number, name what is supposed to become adequate. A stable codebook, a richly textured understanding of a concept, a theme that travels across sites, and enough information to answer a narrow question are not the same target.

Four sources make the distinction concrete. Guest et al. found codebook saturation mostly by 12 interviews in a relatively homogeneous group using a structured guide. Hennink et al. separated code saturation at 9 from meaning saturation for conceptual codes at 16 to 24 in one applied, semi-structured study. Hagaman and Wutich found that common themes in relatively homogeneous sites could appear with 16 or fewer interviews, while cross-site metathemes required roughly 20 to 40 interviews per site in their four-site study. Malterud et al. offer a different planning logic: information power rises with a narrow aim, a specific sample, relevant theory, strong dialogue, and an analysis strategy that does not demand wide comparison.

Table 2 Saturation anchors carry their conditions

Adequacy target	Anchor in the source	Conditions attached to it	Safe use in a new case
Codebook stability	Guest et al.: mostly by 12 interviews	Relatively homogeneous group, structured guide, 60 interviews in the studied dataset	Track new codes and definition changes; do not promise that 12 is enough
Meaning saturation	Hennink et al.: code saturation at 9; meaning saturation at 16 to 24 for conceptual codes	Twenty-five semi-structured interviews in one applied health study; meaning assessed on selected codes	Continue beyond codebook stability when the question needs mechanisms, dimensions, or nuance
Cross-site metathemes	Hagaman and Wutich: roughly 20 to 40 interviews per site	Four cross-cultural sites and a metatheme target, not a single-site codebook	Size each site for the comparison you need; do not import the range into a single corporate sample
Information power	Malterud et al.: five dimensions, no N formula	Aim, sample specificity, theory, dialogue quality, and analysis strategy are appraised together	Make a provisional adequacy judgement, then revisit it as evidence accumulates

Guest, Bunce & Johnson (2006); Hennink, Kaiser & Marconi (2017); Hagaman & Wutich (2017); Malterud, Siersma & Guassora (2016). Each anchor is conditional and cannot be read as a universal N.

For a B2B case with several roles, the practical stop rule is therefore two-layered. First, record whether new conversations add new codes or alter the codebook. Second, record whether they add a new dimension, mechanism, condition, or rival explanation to the focal question. If the sample spans sales, leadership, operations, and a customer, a single pooled count can hide a role that has not been heard. Track the strata that the explanation depends on.

The distinction also protects against a common false precision. A sample of 12 can be too large for a very narrow, information-rich question and too small for a heterogeneous, cross-site comparison. The number is an output of the design, not a substitute for one.

How does rigorous testing of rival explanations confer validity on case studies?

Most case narratives collect support. Stronger case designs collect the prediction of the rival as well. If the claim is that a market-entry result came from local adaptation, a rival may be timing, channel access, a pricing change, or a pre-existing customer relationship. If the claim is that a sales process improved because of a new tool, a rival may be manager attention, territory change, or the removal of weak opportunities.

Write the rival before the evidence is complete. Then keep four fields beside it:

1. What would the rival predict?
2. Which source or observation bears on that prediction?
3. What did the focal explanation predict that the rival did not?

4 What remains ambiguous because the case cannot observe it?

This is why a retrieval audit matters to case work. A correct citation proves that a source exists, not that the source supports the sentence you want. The same distinction applies inside the organisation: a CRM field proves that a value was entered, not that the value caused the result.

A bounded case can explain a mechanism, sequence, or condition. It can show how an outcome became possible and which evidence is consistent with that explanation. It cannot, merely by being detailed, estimate what the average company would experience. A ten-week qualitative case is not weakened by stating this. The boundary is the reason its observations can be interpreted.

How can qualitative and quantitative evidence integrate without methodological pretension?

Adding a dashboard to a case does not automatically make the design mixed methods. Schoonenboom and Johnson describe mixed-methods design as something constructed around purpose, theoretical drive, timing, integration, and priority. The notation QUAL + quan makes a qualitative core and a supplemental quantitative strand visible. If the strands run at different times, the notation and the prose should say so.

The practical implication is demanding but useful. Decide what the quantitative strand is doing: describing the case, checking a sequence, locating a contrast, or testing a proposition. Give it its own data quality and validity conditions. Decide where the strands meet. A table of counts that is never connected to the qualitative explanation is an appendix, not integration.

This also connects to the forecast number you call. A numeric layer can be informative while still being bounded by its cohort, definition, timing, and comparison. Calling it mixed methods does not remove those obligations.

Which diagnostic checklist ensures defensible case study boundaries before fieldwork?

Before the next interview or stakeholder workshop, write down:

1 Question: What decision, mechanism, sequence, or condition is the case meant to illuminate?

2 Unit: Is the unit a company, relationship, project, site, event, team, or decision?

3 Time: What starts the case, what ends it, and which history is only context?

4 Selection: Why this case? What could a contrasting or negative case teach?

5 Evidence: Which source would support the explanation, and which could contradict it?

6 Analysis: What will be compared within the case, across sources, or across cases?

7 Rival: What is the strongest alternative explanation, and what observation would favour it?

8 Transfer: Which conditions may travel, which are local, and which magnitude remains unknown?

9 Stop rule: What new code, dimension, or rival finding would justify one more collection cycle?

If you cannot answer the last three, you have a narrative plan with research language around it. That is fixable. Declare the missing field before the story gets too attractive to question.

Where are the epistemological boundaries of qualitative case study research?

Boundary. This framework improves traceability, rival-explanation discipline, and conditional transfer. It does not turn a purposive case into a random sample or a causal estimate of population- level magnitude. The appropriate stopping point depends on the question, population, evidence quality, and analysis target, and the final transfer claim remains a judgement that must be stated.

Evidence base. The analytical frame also draws on these additional sources: Schoonenboom and Johnson 2017. The links identify the exact works; they support the mechanisms and boundary conditions discussed here, not every claim in isolation.

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

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

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. https://doi.org/10.1177/1525822X05279903

Hagaman, A. K., & Wutich, A. (2017). How many interviews are enough to identify metathemes in multisited and cross-cultural research? Field Methods, 29(1), 23-41. https://doi.org/10.1177/1525822X16640447

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

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

Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 69(S2), 107-131. https://doi.org/10.1007/s11577-017-0454-1

END OF ESSAY

HOW TO CITE THIS ESSAY

Isoglu, S. (2026, August 23). A case study is an evidence design, not a story. Sinan Isoglu. https://isoglu.com/insights/journal/case-study-is-an-evidence-design/

KEEP READING

From the research bench

What is external validity? A result needs a transfer boundary

https://isoglu.com/insights/journal/what-is-external-validity/

From the research bench

Common method bias is not a checkbox

https://isoglu.com/insights/journal/common-method-bias-is-not-a-checkbox/

From the research bench

Mixed reactions are not brand polarization

https://isoglu.com/insights/journal/mixed-reactions-are-not-brand-polarization/

Sinan Isoglu, MBA (Quantic)
Commercial growth leader, lecturer and doctoral researcher

Online version: https://isoglu.com/insights/journal/case-study-is-an-evidence-design/