



FROM THE RESEARCH BENCH

A defensible evidence review has a stopping rule

An evidence review is defensible when its search boundary, inclusion, extraction, rival check, and stopping judgement are visible to readers.

Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher

PUBLISHED

25 August 2026

READING TIME

7 min

WORDS

1,448

<https://isoglu.com/insights/journal/a-defensible-evidence-review-has-a-stopping-rule/>

Management summary	2
Why must literature search boundaries precede keyword query volume?	3
Why is study inclusion criteria an explicit methodological argument?	3
How should researchers document data extraction to ensure auditability?	3
Why is systematic rival hypothesis testing mandatory in evidence reviews?	4
Why does qualitative code saturation fail to signal literature review completion?	4
How should researchers formulate a defensible evidence review stopping rule?	4
Why is evidence review closure an epistemic judgment rather than a count?	5
Where are the methodological boundaries of systematic evidence review rules?	6
References	7

MANAGEMENT SUMMARY

An evidence review becomes a source pile when every new paper is treated as progress and no one can say what would close the search. Eisenhardt describes an iterative process that moves between collection, analysis, comparison, rival explanations, and conditional closure. Gioia makes the extraction structure visible while warning against cookbook use. Biemans, Malshe, and Johnson show a bounded systematic review that codes 73 sales-marketing interface articles through four domains and five themes. Together with conditional saturation and information-power work, these sources support a stopping rule: boundary, inclusion, extraction, disconfirmation, incremental learning, and disclosure.

Keywords: Evidence reviews · Research quality · Literature methods · Decision-making

An evidence review becomes a source pile when every new paper is treated as progress and nobody can say what would close the search.

The answer is not a magic number of sources. It is a visible stopping rule: what is the review object, what is inside it, which extraction structure is being used, what rival could change the answer, and what did the next source add?

Why must literature search boundaries precede keyword query volume?

A review needs a boundary before it needs more results. Name the question, unit, date range, language, source types, and exclusion logic. A search for “customer health” can otherwise expand from churn prediction to support sentiment, retention definitions, product usage, and vendor claims without anyone noticing that the object changed.

Eisenhardt’s account of case-based theory building is useful because it treats collection and analysis as iterative. Questions, case selection, within-case work, cross-case patterns, literature comparison, and closure inform one another. The point is not to search forever. It is to let the search boundary respond to evidence while recording why it changed.

The review record should therefore preserve the opening boundary and every material amendment. A wider search can be legitimate. It should not be hidden as though the original question always meant the final question.

Why is study inclusion criteria an explicit methodological argument?

An inclusion rule is not administrative housekeeping. It defines what can carry the review’s conclusion. Specify the source version, method, population or unit, outcome, and relevant locator before treating a paper as evidence. A source may be useful context and still fail the inclusion rule for a quantitative claim.

The same distinction protects source quality. A working paper can carry a working-paper claim. An abstract can identify a contribution or a gap, but it cannot silently carry a result from an unread full text. A practitioner page can be prior art without becoming evidence for an empirical statement.

The inclusion record should say why a source entered, what it can support, and which claims it cannot carry. That makes the review auditable without pretending that every source has the same evidentiary weight.

How should researchers document data extraction to ensure auditability?

Gioia, Corley, and Hamilton describe an explicit movement from first-order informant terms to second-order concepts and aggregate dimensions. Their approach is a flexible orientation for qualitative rigor, not a cookbook. The transferable lesson is visibility: a reader should see how a source sentence became a code, how codes became a concept, and where the analyst made a judgement.

Systematic reviews need the same discipline in a different form. Biemans, Malshe, and Johnson identify 73 sales-marketing interface articles published from 1990 through 2021. They code the interface through structural linkages, role linkages, interactions, and perceptions, and group the literature into five themes.

That does not make 73 the right size for another review. It shows what a bounded review can make visible: the source universe, the coding domains, the themes, and the relationship between them. The review can be inspected because its structure is part of the result.

Why is systematic rival hypothesis testing mandatory in evidence reviews?

Eisenhardt's process includes pattern searches across cases and comparison with literature. The point is not to produce a supportive table. It is to expose reversals, conditions, and rival explanations. The source that weakens the headline claim may be the most valuable source in the review.

A rival check can ask:

- Does the result disappear under a different unit or denominator?
- Does a source explain the same pattern through a different mechanism?
- Does the method measure a proxy rather than the decision outcome?
- Does the apparent pattern depend on one context, sample, or version?
- What evidence would make the conclusion smaller or false?

If the review only stores confirmatory sources, "saturation" can mean that the search stopped when the argument felt comfortable. That is not a stopping rule. It is an unrecorded preference.

Why does qualitative code saturation fail to signal literature review completion?

Guest, Bunce, and Johnson's 12-interview result is conditional on a relatively homogeneous group and a structured guide. Hennink, Kaiser, and Marconi distinguish code saturation from meaning saturation in one applied semi-structured study. Malterud, Siersma, and Guassora make information power depend on aim, sample specificity, theory, dialogue, and analysis strategy.

These sources are not evidence for a universal interview number, and B12 is not an interview study. They are useful because they separate different closure questions. No new label, no new meaning, no new condition, no new rival, and no new decision implication are not interchangeable observations.

The same distinction applies to a literature review. A source can add no new code and still change a boundary condition. It can repeat a theme and still strengthen a method. It can add no support and still make a causal claim too large. Closure must name which kind of learning has become small.

How should researchers formulate a defensible evidence review stopping rule?

Table 1 Evidence-review stopping rule

Review field	Visible record	Closure question	What remains provisional
Boundary	Question, unit, date, language, source types, exclusions	Is the review object stable enough to compare sources?	Transfer outside the boundary
Inclusion	Version, method, population, outcome, locator, reason for inclusion	Can each source carry the claim assigned to it?	Claims resting only on abstract or context
Extraction	Fields, codes, concepts, themes, and source links	Can a reader follow evidence into the synthesis?	Interpretive judgement
Disconfirmation	Rival, reversal, boundary case, wounded source	What would make the conclusion smaller or false?	Untested rivals
Incremental learning	What the next source changed or failed to change	Has the next source stopped changing answer, conditions, rivals, or uncertainty?	Novel evidence outside the search boundary
Disclosure	Stop date, search record, gaps, versions, and limits	Can another reader understand why closure happened?	Future updates and new source versions

Author's framework grounded in Eisenhardt (1989), Gioia, Corley, and Hamilton (2013), Biemans, Malshe, and Johnson (2022), Guest, Bunce, and Johnson (2006), Hennink, Kaiser, and Marconi (2017), and Malterud, Siersma, and Guassora (2016). This is a review process, not a universal saturation formula.

Why is evidence review closure an epistemic judgment rather than a count?

A review can close while evidence remains incomplete. Closure means that the boundary and the remaining uncertainty are explicit, not that the claim has become universally true. If the next source is unlikely to change the answer within the declared boundary, the source search can stop and the gap can be named.

The opposite is also possible. A review can have hundreds of sources and remain open because the unit, outcome, or rival explanation is unresolved. More papers do not repair a moving question.

The final review record should state the stop date, search boundary, inclusion logic, extraction fields, disconfirmation result, incremental-learning judgement, and unresolved gap. It should also name the source that challenged the argument. That source is part of the evidence review, not an embarrassing footnote.

Where are the methodological boundaries of systematic evidence review rules?

No source above supplies a universal number of papers, interviews, or search hits. Biemans et al. provide one systematic-review example, not a template for every field. Eisenhardt and Gioia provide process and structure, not a guarantee of truth. The contribution is narrower: an evidence review is defensible when its closure can be inspected.

The stopping rule belongs beside case study as evidence design and evidence over anecdote, which make the boundary between observation, synthesis, and transfer visible. It operates alongside the principle that saturation is a decision rule, not a number, defining when further information collection produces diminishing returns.

REFERENCES

- 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>
- Biemans, W., Malshe, A., & Johnson, J. S. (2022). The sales-marketing interface: A systematic literature review and directions for future research. *Industrial Marketing Management*, 102, 324–337. <https://doi.org/10.1016/j.indmarman.2022.02.001>
- 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>
- 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>

END OF ESSAY

HOW TO CITE THIS ESSAY

Isoglu, S. (2026, August 25). A defensible evidence review has a stopping rule. Sinan Isoglu.
<https://isoglu.com/insights/journal/a-defensible-evidence-review-has-a-stopping-rule/>

KEEP READING

From the research bench

Survey bias is a decision error before it is a questionnaire flaw

<https://isoglu.com/insights/journal/survey-bias-is-a-decision-error/>

From the research bench

What is selection bias? The sample can change the answer

<https://isoglu.com/insights/journal/what-is-selection-bias/>

From the research bench

Saturation is a decision rule, not a number

<https://isoglu.com/insights/journal/saturation-is-a-decision-rule-not-a-number/>



Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher