



AUS DER FORSCHUNG

Eine Fallstudie ist ein Evidenzdesign, keine Geschichte.

Ein belastbares Fallstudiendesign setzt Fallgrenzen, prüft Gegen-Erklärungen, plant Sättigung und zeigt, wohin die Evidenz übertragbar ist.

Sinan Isoglu, MBA (Quantic)

Führungskraft für Umsatzwachstum, Dozent und Doktorand

VERÖFFENTLICHT

23. August 2026

LESEZEIT

10 Min.

WÖRTER

2.033

<https://isoglu.com/de/insights/journal/eine-fallstudie-ist-ein-evidenzdesign/>

Management Summary	2
Warum ist ein Unternehmensname nur ein Schauplatz und keine Fallgrenze?	3
Was fordert Eisenhardts Methodik für fundierte qualitative Theoriebildung?	5
Warum lässt sich qualitative Sättigung nicht an starren Interviewzahlen festmachen?	5
Wie verleiht die Prüfung alternativer Hypothesen einer Fallstudie wissenschaftliche Härte?	6
Wie lassen sich qualitative und quantitative Befunde ohne methodischen Ballast verknüpfen?	7
Welche Checkliste sichert belastbare Fallgrenzen vor dem Start der Datenerhebung?	7
Wo liegen die epistemologischen Grenzen qualitativer Fallstudienforschung?	8
Quellen	9

MANAGEMENT SUMMARY

Eine Fallstudie wird nicht durch ein bekanntes Unternehmen, einige Interviews oder viele Quellen belastbar. Sie wird es durch ein offengelegtes Design: Frage, Fallgrenze, Untersuchungseinheit, Zeitraum, Auswahl, Quellen, Analyse, Gegen-Erklärungen und Bedingungen der Übertragbarkeit. Dieser Beitrag verbindet Eisenhardts iterativen Prozess mit der Gioia-Datenstruktur und ordnet vier bedingte Ansätze zur Angemessenheit der Stichprobe ein. Die zwölf Interviews bei Guest, die Trennung von Code- und Bedeutungssättigung bei Hennink, die standortübergreifenden Metathemen bei Hagaman und Wutich sowie Malteruds Informationsstärke sind Planungsanker, keine allgemeingültigen Fallzahlen. Die Checkliste zeigt, was vor dem nächsten Gespräch feststehen sollte. Ein purposiver Fall kann einen Mechanismus oder eine Bedingung erklären, aber keine Kausalstärke für die Grundgesamtheit liefern.

Schlagwörter: Fallstudienforschung · Forschungsdesign · Qualitative Forschung · Evidenzqualität · Mixed Methods

Ein Kundenlogo, acht Interviews und eine lange Literaturliste sehen schnell nach Forschung aus. Die entscheidende Frage bleibt aber offen: Was genau ist hier der Fall, und welche Beobachtung würde der Erklärung widersprechen?

Eine Fallstudie ist deshalb kein besonders anschaulicher Bericht. Sie ist ein begrenztes Design, das eine Frage mit nachvollziehbarer Evidenz beantwortet. Dazu gehören die Fallgrenze, die Einheit, der Zeitraum, die Auswahl, die Quellen, die Analyse, die Gegen-Erklärungen und die Aussage darüber, was übertragbar ist. Das ist die gleiche Grundhaltung wie bei der Evidenzschwelle: Nicht das Dokument adelt die Zahl, sondern der Prüfweg, den sie überstanden hat.

Warum ist ein Unternehmensname nur ein Schauplatz und keine Fallgrenze?

Ein Unternehmen kann der Kontext sein, ohne die Untersuchungseinheit zu sein. Bei einem Markteintritt kann der Fall die Entscheidung, die ersten drei Länder, die Beziehung zum Vertriebspartner oder die Abfolge eines Konflikts sein. Jede dieser Grenzen verändert die Quellen, die Gegenbeispiele und die Aussage, die am Ende zulässig ist.

Die Grenze gehört deshalb vor die interessanteste Geschichte. Schreiben Sie auf, was in den Fall gehört, was außerhalb bleibt, wann die Beobachtung beginnt und endet und welchen Mechanismus oder welches Ergebnis die Untersuchung beleuchten soll. Die folgende Matrix macht aus diesen Fragen ein Arbeitsobjekt. Links steht, was ein Erzählbericht häufig offenlässt. Rechts steht, was ein Evidenzdesign sichtbar macht.

Table 1 Vom Erzählfall zum Evidenzdesign

Designfeld	Erzählversion	Evidenzversion
Frage	Was ist in diesem Unternehmen passiert?	Welche Entscheidung, welcher Mechanismus oder welche Bedingung wird untersucht?
Fallgrenze	Name des Unternehmens, der Marke oder des Projekts	Explizite Ein- und Ausschlussregeln mit Anfang und Ende
Einheit	Die Organisation als einheitlicher Akteur	Entscheidung, Beziehung, Ereignis, Team, Standort oder Organisation, die zur Frage passt
Auswahl	Der sichtbarste Erfolg oder der einfachste Zugang	Ein theoretischer Grund: Replikation, Erweiterung, Kategorie oder Kontrastfall
Evidenz	Interviews, die die Chronologie lesbar machen	Interviews, Beobachtung, Dokumente, Systemspuren oder Messungen mit Herkunftsnachweis
Analyse	Glatte Chronologie und Themenliste	Rekonstruktion innerhalb des Falls, Vergleich von Quellen und Suche nach Umkehrungen
Gegen-Erklärung	Eine Fußnote mit einer Alternative	Log der Vorhersage, der stützenden Beobachtungen und der offenen Punkte
Übertragbarkeit	So funktioniert es	Diese Bedingungen können übertragbar sein; diese bleiben lokal; diese Größe ist nicht identifiziert

Eigene Operationalisierung nach Eisenhardt (1989) und Gioia, Corley & Hamilton (2013). Die rechte Spalte ist ein Arbeitsinstrument, kein veröffentlichtes Messinstrument.

Die schwierigste Zeile ist oft die Auswahl. Ein erfolgreicher Fall ist nicht automatisch falsch gewählt. Problematisch wird er, wenn die Auswahl stillschweigend als typisches Ergebnis gelesen wird. Theoretisches Sampling gibt der Auswahl einen Lernzweck: Ein Fall kann ein Muster replizieren, auf eine andere Bedingung erweitern, eine begriffliche Kategorie füllen oder einen Kontrast sichtbar machen. Das erzeugt keine statistische Repräsentativität. Es sagt, was dieser Fall zeigen soll.

Auch die Quellen sollten sich widersprechen dürfen. Ein Kundengespräch kann eine Wahrnehmung erklären. Eine CRM-Spur kann eine Abfolge zeigen. Ein Vertrag kann klären, was vereinbart war. Ein verlorenes Angebot oder ein widersprüchliches internes Dokument kann eine elegante Erklärung stärker belasten als ein weiteres zustimmendes Gespräch. Triangulation bedeutet daher nicht, möglichst viele gleiche Stimmen zu sammeln. Sie bedeutet, Quellen zusammenzustellen, die unterschiedliche Fehlerstellen haben.

Was fordert Eisenhardts Methodik für fundierte qualitative Theoriebildung?

Eisenhardts Beitrag von 1989 ist für diese Unterscheidung nützlich, weil er Fallstudienforschung als Bewegung zwischen Frage und Evidenz beschreibt. Der Prozess beginnt mit einer breiten Frage und einer vorläufigen begrifflichen Orientierung. Danach folgen theoretische Fallauswahl, mehrere Methoden oder Forschende, die Überlappung von Datenerhebung und Analyse, Vertrautheit mit dem Einzelfall, Vergleiche über Fälle oder Quellen hinweg, Literaturvergleich und eine begründete Schließung, wenn der zusätzliche theoretische Erkenntnisgewinn klein wird.

Das verändert auch die kommerzielle Fallarbeit. Sie warten nicht bis zum Ende, um herauszufinden, was die Evidenz bedeutet. Sie formulieren die erste Erklärung, sammeln die Beobachtungen, die sie vorhersagt, sucht gezielt nach einer möglichen Umkehrung und passt die Grenze an, wenn sich die Frage verändert. Die Anpassung ist kein Mangel, wenn sie dokumentiert wird. Eine unbemerkte Anpassung macht aus einem Untersuchungsfall schnell eine Geschichte über das, was das Team bereits glaubt.

Gioia, Corley und Hamilton ergänzen eine sichtbare Strukturebene. First-Order-Codes halten die Begriffe der Befragten nahe am Material. Second-Order-Konzepte bewegen sich in die theoretische Sprache der Analyse. Aggregate Dimensionen zeigen die übergeordnete Struktur. So kann eine Leserin nachvollziehen, wie die Schlussfolgerung gebaut wurde.

Die Warnung ist ebenso wichtig wie die Struktur: Gioia et al. verstehen den Ansatz als flexible Orientierung, nicht als Kochbuch. Ein Rahmen darf nicht zum Abzeichnen werden. Wenn das Material keine dreistufige Struktur trägt, erzwingen Sie sie nicht. Wenn Systemspuren und Gespräche einander widersprechen, ist der Widerspruch nicht automatisch ein Datenfehler.

Warum lässt sich qualitative Sättigung nicht an starren Interviewzahlen festmachen?

Die Frage nach der nötigen Interviewzahl wird meist zu spät gestellt. Zuerst muss feststehen, was ausreichend werden soll: ein stabiles Codebuch, ein nuanciertes Verständnis eines Konzepts, ein Thema über mehrere Standorte hinweg oder genügend Information für eine enge Frage. Das sind verschiedene Ziele.

Guest et al. fanden in einer relativ homogenen Gruppe mit strukturiertem Leitfaden, dass die Codebuchsättigung in ihrem Datensatz überwiegend bis Interview 12 erreicht war. Hennink et al. trennten in einer anwendungsbezogenen Studie die Codesättigung bei 9 von der Bedeutungssättigung konzeptioneller Codes bei 16 bis 24 Interviews. Hagaman und Wutich zeigen den Unterschied zwischen gemeinsamen Themen innerhalb relativ homogener Standorte und standortübergreifenden Metathemen. Für Letztere brauchte ihre Vier-Standorte-Studie ungefähr 20 bis 40 Interviews je Standort. Malterud et al. schlagen eine andere Planungssprache vor: Informationsstärke steigt mit enger Frage, spezifischer Stichprobe, passender Theorie, guter Gesprächsqualität und einer Analyse-methode, die keinen weiten Vergleich verlangt.

Table 2 Sättigungsanker tragen ihre Bedingungen

Ziel der Angemessenheit	Anker in der Quelle	Bedingungen	Sichere Nutzung im neuen Fall
Codebuchstabilität	Guest et al.: überwiegend bis Interview 12	Relativ homogene Gruppe, strukturierter Leitfaden, 60 Interviews im Datensatz	Neue Codes und Definitionsänderungen verfolgen; 12 nicht versprechen
Bedeutungssättigung	Hennink et al.: Codes bei 9, Bedeutung konzeptioneller Codes bei 16 bis 24	25 semistrukturierte Gespräche in einer anwendungsbezogenen Gesundheitsstudie	Nach stabilen Codes weiterarbeiten, wenn Mechanismen, Dimensionen oder Nuancen wichtig sind
Standortübergreifende Metathemen	Hagaman und Wutich: ungefähr 20 bis 40 Interviews je Standort	Vier kulturübergreifende Standorte und ein Metathemenziel, kein einzelnes Codebuch	Jeden Standort für den geplanten Vergleich bemessen; den Bereich nicht in eine Unternehmensstichprobe übertragen
Informationsstärke	Malterud et al.: fünf Dimensionen, keine Formel	Ziel, Spezifität, Theorie, Gesprächsqualität und Analysemethoden gemeinsam beurteilen	Eine vorläufige Einschätzung treffen und während der Erhebung neu prüfen

Guest, Bunce & Johnson (2006); Hennink, Kaiser & Marconi (2017); Hagaman & Wutich (2017); Malterud, Siersma & Guassora (2016). Jeder Anker ist bedingt und keine allgemeingültige Fallzahl.

Für einen B2B-Fall mit mehreren Rollen braucht die Stop-Regel deshalb zwei Ebenen. Erstens wird festgehalten, ob neue Gespräche neue Codes oder Änderungen am Codebuch bringen. Zweitens wird geprüft, ob eine neue Dimension, ein Mechanismus, eine Bedingung oder eine Gegen-Erklärung hinzukommt. Wenn Vertrieb, Führung, Operations und Kundenseite vorkommen, kann eine gemeinsame Zahl eine Rolle verdecken, die noch nicht ausreichend gehört wurde.

Wie verleiht die Prüfung alternativer Hypothesen einer Fallstudie wissenschaftliche Härte?

Erzählberichte sammeln meistens Bestätigung. Ein belastbares Design sammelt auch die Vorhersage der Gegen-Erklärung. Wenn das Ergebnis eines Markteintritts mit lokaler Anpassung erklärt wird, können Zeitpunkt, Vertriebspartner, Preisänderung oder eine bestehende Kundenbeziehung die Alternative sein. Wenn ein neues Tool einen Vertriebsprozess verbessert haben soll, können Managementaufmerksamkeit, Territorien oder das Aussortieren schwacher Chancen die Gegen-Erklärung bilden.

Formulieren Sie die Alternative, bevor die Evidenz vollständig ist. Halten Sie daneben vier Felder fest:

- 1 Was sagt die Gegen-Erklärung voraus?
- 2 Welche Quelle oder Beobachtung prüft diese Vorhersage?
- 3 Was sagte die bevorzugte Erklärung voraus, die die Gegen-Erklärung nicht erklärt?
- 4 Was bleibt unklar, weil der Fall es nicht beobachten kann?

Darum ist auch eine Retrieval-Prüfung für Fallarbeit wichtig. Eine korrekte Quellenangabe zeigt, dass ein Text existiert. Sie zeigt nicht automatisch, dass er den gewünschten Satz trägt. Im Unternehmen gilt dasselbe: Ein CRM-Feld zeigt, dass ein Wert eingetragen wurde, nicht dass er das Ergebnis verursacht hat.

Ein begrenzter Fall kann einen Mechanismus, eine Abfolge oder eine Bedingung erklären. Er kann zeigen, wie ein Ergebnis möglich wurde und welche Evidenz zu dieser Erklärung passt. Durch Detailreichtum wird er aber nicht zur Schätzung dessen, was ein durchschnittliches Unternehmen erleben würde. Eine zehnwöchige qualitative Fallstudie wird durch diese Grenze nicht schwächer. Erst die Grenze macht die Beobachtungen interpretierbar.

Wie lassen sich qualitative und quantitative Befunde ohne methodischen Ballast verknüpfen?

Ein Dashboard neben einer Fallbeschreibung macht das Design noch nicht zu Mixed Methods. Schoonenboom und Johnson beschreiben Mixed-Methods-Designs als Konstruktion aus Zweck, theoretischer Orientierung, zeitlicher Abfolge, Integration und Priorität. Die Notation QUAL + quan macht sichtbar, dass ein qualitativer Kern von einer quantitativen Ergänzung begleitet wird. Wenn die Stränge zu verschiedenen Zeitpunkten laufen, muss auch das so benannt werden.

Die praktische Konsequenz ist anspruchsvoll. Entscheiden Sie, ob der quantitative Strang den Fall beschreibt, eine Abfolge prüft, einen Kontrast lokalisiert oder eine Proposition testet. Legen Sie seine Qualitäts- und Validitätsbedingungen separat fest. Entscheiden Sie, an welchem Punkt die Stränge zusammengeführt werden. Eine Zähltablette ohne Verbindung zur qualitativen Erklärung ist ein Anhang, keine Integration.

Das gilt auch für die Forecast-Zahl, die gemeldet wird. Eine numerische Ebene kann informativ sein und trotzdem durch Kohorte, Definition, Zeitpunkt und Vergleich begrenzt bleiben. Das Etikett Mixed Methods nimmt diese Pflichten nicht weg.

Welche Checkliste sichert belastbare Fallgrenzen vor dem Start der Datenerhebung?

Vor dem nächsten Interview oder Workshop sollten neun Fragen schriftlich beantwortet sein:

- 1 Frage: Welche Entscheidung, welcher Mechanismus, welche Abfolge oder Bedingung soll sichtbar werden?
- 2 Einheit: Ist die Einheit ein Unternehmen, eine Beziehung, ein Projekt, ein Standort, ein Ereignis, ein Team oder eine Entscheidung?
- 3 Zeit: Was beginnt den Fall, was beendet ihn, und welche Historie ist nur Kontext?
- 4 Auswahl: Warum dieser Fall? Was würde ein Kontrast- oder Negativfall zeigen?
- 5 Evidenz: Welche Quelle stützt die Erklärung, und welche könnte ihr widersprechen?
- 6 Analyse: Was wird innerhalb des Falls, zwischen Quellen oder zwischen Fällen verglichen?
- 7 Gegen-Erklärung: Was ist die stärkste Alternative, und welche Beobachtung würde sie stützen?
- 8 Übertragbarkeit: Welche Bedingungen können wandern, welche bleiben lokal, und welche Größe bleibt unbekannt?
- 9 Stop-Regel: Welcher neue Code, welche Dimension oder Gegen-Erklärung rechtfertigt eine weitere Erhebungsrunde?

Wenn die letzten drei Fragen offen sind, liegt ein Erzählplan mit Forschungssprache vor. Das lässt sich reparieren. Die fehlende Entscheidung muss sichtbar werden, bevor die Geschichte zu attraktiv wird, um sie noch zu prüfen.

Wo liegen die epistemologischen Grenzen qualitativer Fallstudienforschung?

Grenze. Der Rahmen verbessert Nachvollziehbarkeit, den Umgang mit Gegen-Erklärungen und die bedingte Übertragbarkeit. Es macht aus einem purposiven Fall weder eine Zufallsstichprobe noch eine Kausalschätzung für die Grundgesamtheit. Der angemessene Abbruch hängt von Frage, Population, Evidenzqualität und Analyseziel ab. Die abschließende Übertragung bleibt ein Urteil und muss als solches benannt werden.

Evidenzbasis. Der analytische Rahmen stützt sich außerdem auf diese zusätzlichen Quellen: Schoonenboom and Johnson 2017. Die Links identifizieren die konkreten Werke; sie stützen die hier diskutierten Mechanismen und Geltungsgrenzen, nicht jede einzelne Aussage isoliert.

- 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>

ENDE DES BEITRAGS

ZITIERWEISE

Isoglu, S. (2026, 23. August). Eine Fallstudie ist ein Evidenzdesign, keine Geschichte. Sinan Isoglu. <https://isoglu.com/de/insights/journal/eine-fallstudie-ist-ein-evidenzdesign/>

WEITERLESEN

Aus der Forschung

Gemischte Reaktionen sind keine Markenpolarisierung

<https://isoglu.com/de/insights/journal/gemischte-reaktionen-sind-keine-markenpolarisierung/>

Aus der Forschung

Was ist externe Validität? Ein Ergebnis braucht eine Transfergrenze

<https://isoglu.com/de/insights/journal/was-ist-externe-validitaet/>

Aus der Forschung

Ein öffentliches Forum ist keine Marktbefragung

<https://isoglu.com/de/insights/journal/ein-oeffentliches-forum-ist-keine-marktbefragung/>



Sinan Isoglu, MBA (Quantic)

Führungskraft für Umsatzwachstum, Dozent und Doktorand