Theory-Building With Cases

Tony Hak and Jan Dul

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2009-036-ORG
Publication	June 2009
Number of pages	12
Persistent paper URL	http://hdl.handle.net/1765/16205
Email address corresponding author	thak@rsm.nl; jdul@rsm.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus Universiteit Rotterdam P.O.Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

ERASMUS **R**ESEARCH **I**NSTITUTE OF **M**ANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

ABSTRACT AND KEYWORDS	
Abstract	Theory-building with cases is (a) formulating new propositions that emerge from the empirical evidence in a sample of cases and (b) testing them in the same sample. The main difference with most other forms of generating new propositions (such as analyzing the theoretical literature, brainstorming, etc.) is its empirical character. The main difference with other forms of discovering new propositions in empirical evidence (such as in 'exploratory' research) is that only those theoretical formulations are accepted as a result of the theory-building study that are confirmed in a test in the sample from which the proposition was built. It is possible that a proposition about a relationship between two variables emerges from an exploratory single case study (e.g., when both variables have extreme values in that case), but it is not possible to test that new proposition in the same study because this would require a comparison in a sample of cases. The term theory-building study (as distinct from an exploratory study) is used here only for studies in which a proper test of the new proposition has been conducted.
Free Keywords	theory-building; case study research; necessary condition; sufficient condition
Availability	
Classifications	

Theory-Building With Cases

Tony Hak & Jan Dul

Rotterdam School of Management

thak@rsm.nl; jdul@rsm.nl

Entry for the Encyclopedia of Case Study Research,

edited by Albert J. Mills, Gabrielle Durepos, & Elden Wiebe

(Sage, 2009)

Final Version October 2008

Keywords: Theory-building; Case study research; Necessary condition; Sufficient condition

Abstract

Theory-building with cases is (a) formulating new propositions that emerge from the empirical evidence in a sample of cases and (b) testing them in the same sample. The main difference with most other forms of generating new propositions (such as analyzing the theoretical literature, brainstorming, etc.) is its empirical character. The main difference with other forms of discovering new propositions in empirical evidence (such as in 'exploratory' research) is that only those theoretical formulations are accepted as a result of the theory-building study that are confirmed in a test in the sample from which the proposition was built. It is possible that a proposition about a relationship between two variables emerges from an exploratory single case study (e.g., when both variables have extreme values in that case), but it is not possible to test that new proposition in the same study because this would require a comparison in a sample of cases. The term theory-building study (as distinct from an exploratory study) is used here only for studies in which a proper test of the new proposition has been conducted.

Conceptual Overview and Discussion

A theory-building study always requires a sample of cases, because multiple cases are needed for the test of the new propositions. In the literature there is only one influential and frequently cited approach to theory-building with cases, the Yin-Eisenhardt approach. First this approach will be described and evaluated. Then a revised approach will be discussed and illustrated. For illustrative purposes an example is used that is taken from another publication by Kathleen Eisenhardt in which she builds a number of new propositions concerning the determinants and effects of fast strategic decision making in firms in high-velocity environments based on empirical evidence collected in eight microcomputer firms. The example used here is the proposition "*The greater the use of real-time information, the greater the speed of the strategic decision process*".

Applications

The Yin-Eisenhardt Approach To Theory-Building With Cases

The core of the Yin-Eisenhardt approach to theory-building with cases consists of three steps: (a) within-case analysis, (b) cross-case pattern search, (c) replication.

(a) ***Within-case analysis***

Within-case analysis typically involves detailed case study write-ups for each site. These write-ups are often simply pure descriptions. The overall idea is to become intimately familiar with each case as a stand-alone entity. As the unit of analysis in the study of decision making in microcomputer firms is the decision, the case study write-up in this study was a 'decision story' which was developed by combining the collected data into a time line that included all events relevant to the decision.

(b) ***Cross-case pattern search***

A number of 'tactics' can be used in this crucial step in which the theory (at least its building blocks, propositions) is actually built. One tactic is to select categories or dimensions, and then to look for within-group similarities coupled with intergroup differences. A second tactic is to select pairs of cases and then to list the similarities and differences between each pair. Overall, the idea behind these (and other) cross-case searching tactics is to force investigators to go beyond initial impressions. Consider the following example of a pair of cases (Table 1).

Table 1

Firm	Vice President for Finance?	Number of Routine Quantative Targets	Number of Weekly Operations Meetings	Decision Duration in Months
Zap	Yes	6+	3	3
Presidential	No	3	0	18

Real-time information and decision duration. Two cases

Source: Eisenhardt (1989b), Tables 2 and 3

Starting from the evidence in the 'decision stories' generated in the 'within-case analysis', a number of indicators of real-time information use were developed, such as (a) the presence of a vice president for finance, (b) the number and kind of performance measures and targets that are reviewed regularly, (c) the number of meetings regularly scheduled to review current operations. Table 1 presents the scores on these indicators as well as the decision duration for two cases. It is not difficult to see how a proposition such as "*The greater the use of real-time information, the greater the speed of the strategic decision process*" could 'emerge' from the evidence in this pair of cases and how the researcher's belief in this proposition could be reinforced by evidence from other pairs of cases showing a similar pattern.

(c) ***Replication***

In the Yin-Eisenhardt approach it is considered necessary to verify that the emergent propositions fit with the evidence in each case. This verification process is described as similar to that in traditional hypothesis testing research. The key difference is that each hypothesis is examined for each case, not for the aggregate cases. Thus, the underlying logic is replication, that is, the logic of treating a series of cases as a series of experiments with each case serving to confirm or disconfirm the hypotheses. Replication is the core of the Yin-Eisenhardt approach and hence most studies that claim to have applied this approach state that replication took place. There is, however, no published example of an actual case of replication. A probable reason for this absence is the fact that often a replication in a single case is not possible. Consider again the proposition "*The greater the use of real-time information, the greater the speed of the strategic decision process*". This proposition states that a *change* or a *difference* in the use of real-time information is associated with a *change* or a *difference* in decision speed. The correctness of this proposition cannot be examined for each separate case because it is not a proposition about the characteristics of single cases. The underlying problem is that no distinction is made between (a) propositions about characteristics of single cases, and (b) propositions about differences between cases. The Yin-Eisenhardt approach needs to be amended for this reason. The alternative approach, discussed below, retains the very important idea that theory-building must entail a final, third, step of testing. The main difference with the Yin-Eisenhardt approach is that it distinguishes different types of proposition from the outset.

Revised Approach To Theory-Building With Cases

The revised approach to theory-building with cases entails the same three steps as in the Yin-Eisenhardt approach: (a) within-case analysis, (b) cross-case pattern search, (c) testing.

(a) ***Within-case analysis***

Within-case analysis boils down to what is called 'measurement' in other research strategies. Evidence is collected about relevant variables and this evidence is transformed into scores that indicate relevant levels of the values of these variables, so scores can be compared between cases. These scores can be presented in a data matrix (as in Table 2). A data matrix is a list of scores of the independent and dependent variables for each single case. These data can be qualitative or quantitative, and can be obtained by collecting new data, or by extracting them from existing databases.

Table 2

Firm	Vice President for Finance?	Number of Routine Quantative Targets	Number of Weekly Operations Meetings	Decision Duration in Months
Triumph	Yes	5+	3	1.5
Forefront	Yes	5+	2	2
Zap	Yes	6+	3	3
Promise	Yes	4+	2	4
Omicron	Yes, but weak	None observed	2	6
Neutron	Yes	None observed	1	12
Alpha	Yes	2	1	12
Presidential	No	3	0	18

Real-time information and decision duration

Source: Eisenhardt (1989b), Tables 2 and 3

(b) ***Cross-case pattern search***

It is recommended to start the process of discovering relations in a data matrix with determining whether there is evidence for sufficient conditions (i.e., causes that automatically result in an outcome), next for necessary conditions (i.e., causes that must be present for an outcome to occur), and finally for relations between changes or differences in the values of variables. Methods that can be used for each of these search tactics will be described, and some of them will be illustrated with recoded data as presented in Table 3 (which is a recoded version of Table 2). This recoding, which requires criteria (not discussed here), is necessary for the discovery of sufficient conditions and of necessary conditions with *discrete* variables (e.g., variables with only two scores such as ‘present’ and ‘absent’, or with a limited set of possible values such as ‘high’, ‘medium’ and ‘low’).

Table 3

Firm	Vice President for Finance?	Number of Routine Quantative Targets	Number of Weekly Operations Meetings	Decision Speed
Triumph	Yes	Many	3	Fast
Forefront	Yes	Many	2	Fast
Zap	Yes	Many	3	Fast
Promise	Yes	Many	2	Fast
Omicron	Yes, but weak	None observed	2	Medium
Neutron	Yes	None observed	1	Slow
Alpha	Yes	Few	1	Slow
Presidential	No	Few	0	Slow

Real-time information and decision duration

Source: Eisenhardt (1989b), Tables 2 and 3

Table 4

Number of Routine Quantative Targets	Decision Speed
Many	**Fast**
Many	**Fast**
Many	**Fast**
Many	**Fast**
None observed	Medium
None observed	Slow
Few	Slow
Few	Slow

(1) ***Sufficient condition***

A sufficient condition proposition states that a specific value of a causal variable *always* results in a specific outcome (i.e., a specific value of a dependent variable). The method of finding such conditions in a data matrix consists of ascertaining for each value of an independent variable that occurs more than once whether it always is related to the same outcome.

Table 4 (assuming "none observed" as zero for Number of Routine Quantitative Targets) shows that having Many (i.e., more than four) targets (occurring 4 times in the data matrix), is always associated with Fast decisions. Hence the first candidate proposition that is supported by the data matrix is:

- *Having many (>4) routine quantitative targets is a sufficient condition for fast (≤4 month) decisions.*

Table 5

Number of Weekly Operations Meetings	Decision Speed
3	**Fast**
2	Fast
3	**Fast**
2	Fast
2	Medium
1	Slow
1	Slow
0	Slow

Table 5 shows that having three weekly operations meetings (a score that occurs two times in the data matrix) is always associated with decision speeds of 4 months or less. Hence another candidate proposition that is supported by the data matrix is:

- *Having three weekly operations meetings is a sufficient condition for fast (≤4 month) decisions.*

(2) ***Necessary condition***

A necessary condition proposition states that an outcome can only exist if a specific value of a causal variable exists. The method of finding such conditions in a data matrix consists of identifying values of the independent variables that always occur if the outcome is present. In the present example, therefore, only cases in which the outcome occurs are used (see Table 6) and in these cases those values of the condition are identified that occur in all four cases. Three necessary conditions can be identified:

- *Having a VP for finance is a necessary condition for fast decisions.*
- *Having many (=4+) routine quantitative targets is a necessary condition for fast decisions.*
- *Having at least two weekly operations meetings is a necessary condition for fast decisions.*

Table 6

Firm	Vice President for Finance?	Number of Routine Quantative Targets	Number of weekly Operations Meetings	Decision Speed
Triumph	Yes	Many	3	Fast
Forefront	Yes	Many	2	Fast
Zap	Yes	Many	3	Fast
Promise	Yes	Many	2	Fast

Combining the findings regarding necessary and sufficient conditions, a proposition can be formulated about a condition that is both necessary *and* sufficient:

- *Having many (=4+) routine quantitative targets is a necessary and sufficient condition for fast decisions.*

(3) ***Relations***

The main method for discovering continuously increasing or decreasing (linear or more complex) relations is comparing rankings. In Tables 7 and 8 the cases are ranked according to their number of routine quantitative targets (Table 7) and their number of weekly operations meetings (Table 8) respectively. For these rankings the original scores from Table 2 are used.

Table 7

Number of Routine Quantative Targets	Decision Duration in Months
6+	3
5+	1.5
5+	2
4+	4
3	18
2	12
None observed	6
None observed	12

The relation that explains the data in Table 7 could be expressed in the following proposition:

- *The higher the number of routine quantitative targets, the faster the decision.*

Table 7 does not show a perfect increase in the decision duration column. A scatter plot would show two quite separate clusters of four cases each (if "none observed" is recoded as zero). The proposition above in which the relation is formulated as a necessary and sufficient condition (*"Having many routine quantitative targets is a necessary and sufficient condition for fast decisions"*) seems a better expression of the relation between the values in the two columns in Table 7.

Table 8

Number of Weekly Operations Meetings	Decision Duration in Months
3	1.5
3	3
2	2
2	4
2	6
1	12
1	12
0	18

In Table 8 an almost perfect increase of scores can be observed in the column of the dependent variable, which is a strong ground for building the following proposition:

- *The higher the number of weekly operations meetings, the faster the decision.*

By applying the revised approach to the same data set from which Eisenhardt developed the proposition that is taken here as an example, a number of candidate propositions have been built. Not all are equally useful for the theory. For example, necessary condition propositions can be trivial if the condition is always present. Because all but one firm in this data set, and arguably almost all firms in the theoretical domain, have a VP for finance, the necessary condition proposition regarding the VP for finance might be considered trivial. It is possible that only one of the three necessary conditions will appear to be a 'real' necessary condition, and that the other two just happened to co-occur with that condition in this data set.

If looked at the list of candidate propositions in this way, two can be selected as likely the most important:

- *Having many (=4+) routine quantitative targets is a necessary and sufficient condition for fast decisions.*
- *The higher the number of weekly operations meetings, the faster the decision.*

These two propositions are complementary. The first one formulates a necessary condition for fast decisions which also seems to be sufficient for fastness. The latter one explains or predicts additional variation in speed within subgroups of fast or slow decision speed. This is a more specific result than was generated by Eisenhardt from the same data. It demonstrates that it is useful to apply the revised methods described here to build propositions.

(c) ***Testing***

The propositions that have been built in the cross-case search now must be tested in the data set. Eisenhardt did not build a necessary or sufficient condition proposition. Interestingly, the method of replication as formulated as part of the Yin-Eisenhardt approach (with each case serving to confirm or disconfirm the hypotheses), which is not applicable in building a proposition about relations between differences in values of variables, actually can be applied in the building of a necessary or sufficient condition proposition. A replication would consist of applying the methodology of the theory-testing single case study to each case of the data set.

If the two methodologies described in the entry on Theory-testing with cases are applied on the two propositions built from Table 2, no rejections will be found.

Critical Summary

In sum, the revised approach to theory-building with cases as presented here entails the same three steps as the Yin-Eisenhardt approach. In the first step (‘within-case analysis’) the data matrix with scores of the independent and dependent variables for each case is built. These scores can be quantitative as well as qualitative, and can be obtained by collecting new data, or by extracting them from existing databases. In the second step (‘cross-case pattern search’) different tactics are applied for discovering different kinds of proposition: sufficient conditions, necessary conditions, and relations. In the third step (‘testing’), the methodology for theory-testing with cases (as discussed in the entry Theory-testing with cases) is applied to confirm the fact that the propositions have been built properly.

Further reading and references

Dul, J. & Hak, T. (2008). *Case Study Research Methodology in Business Research*. Oxford: Butterworth-Heinemann.

Eisenhardt, K.M. (1989a), Building theories from case study research, *Academy of Management Review*, 14(4), 532-550.

Eisenhardt, K.M. (1989b), Making fast strategic decisions in high-velocity environments, *Academy of Management Journal*, 12(3), 543-576.

Yin, R.K. (1984), *Case Study Research. Design and Methods*, Sage, Thousand Oaks, CA.

Publications in the ERIM Report Series Research* in Management

ERIM Research Program: "Organizing for Performance"

2009

On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-Inflated Estimation
Martijn J. Burger, Frank G. van Oort, and Gert-Jan M. Linders
ERS-2009-003-ORG
http://hdl.handle.net/1765/14614

The Current State of Research on Networks in China's Business System
Johannes Meuer and Barbara Krug
ERS-2009-007-ORG
http://hdl.handle.net/1765/14842

Encountered Problems and Outcome Status in Nascent Entrepreneurship
A. Marco van Gelderen, B. Pankaj Patel, and A. Roy Thurik
ERS-2009-008-ORG
http://hdl.handle.net/1765/14913

The Ethics of Organizations: A Longitudinal Study of the U.S. Working Population
Muel Kaptein
ERS-2009-018-ORG
http://hdl.handle.net/1765/15405

From Symbolic to Substantive Documents: When Business Codes of Ethics Impact Unethical Behavior in the Workplace
Muel Kaptein
ERS-2009-021-ORG
http://hdl.handle.net/1765/15909

Knowledge Management as a Strategic Tool to Foster Innovativeness of SMEs
Haibo Zhou, and Lorraine Uhlaner
ERS-2009-025-ORG
http://hdl.handle.net/1765/15913

Knowledge Management in the SME and its Relationship to Strategy, Family Orientation and Organization Learning
Haibo Zhou, and Lorraine Uhlaner
ERS-2009-026-ORG
http://hdl.handle.net/1765/15914

Defining Respectful Leadership: What it is, how it can be measured, and another glimpse at what it is related to
Niels van Quaquebeke, and Tilman Eckloff
ERS-2009-027-ORG
http://hdl.handle.net/1765/15942

Blue Ocean versus Competitive Strategy: Theory and Evidence
Andrew Burke, André van Stel, and Roy Thurik
ERS-2009-030-ORG
http://hdl.handle.net/1765/16037

Pattern matching
Tony Hak and Jan Dul
ERS-2009-034-ORG
http://hdl.handle.net/1765/16203

Replication
Tony Hak and Jan Dul
ERS-2009-035-ORG
http://hdl.handle.net/1765/16204

Theory-Building With Cases
Tony Hak and Jan Dul
ERS-2009-036-ORG
http://hdl.handle.net/1765/16205

Theory-Testing With Cases
Tony Hak and Jan Dul
ERS-2009-036-ORG
http://hdl.handle.net/1765/16206

∗ A complete overview of the ERIM Report Series Research in Management:
https://ep.eur.nl/handle/1765/1

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship