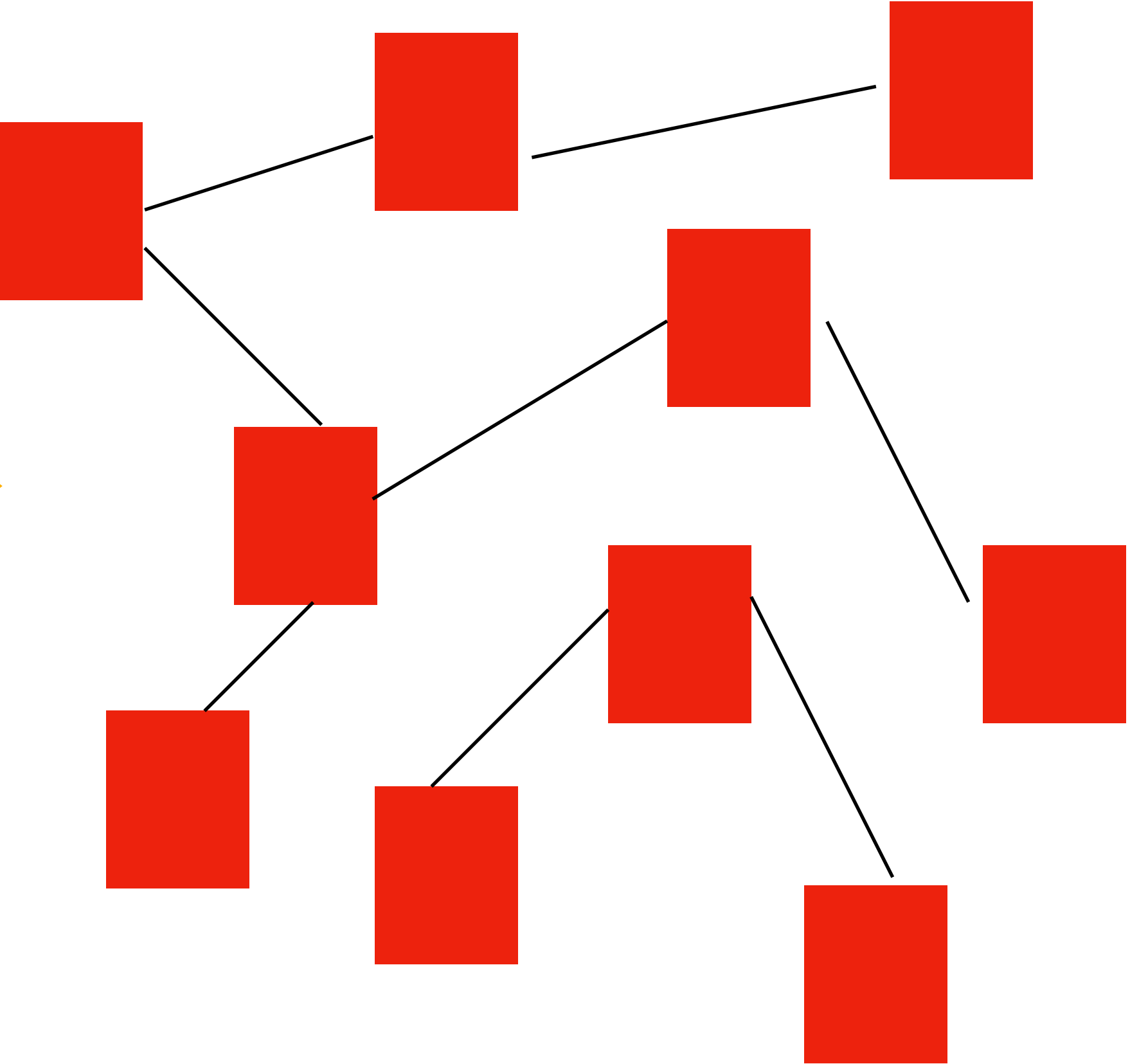
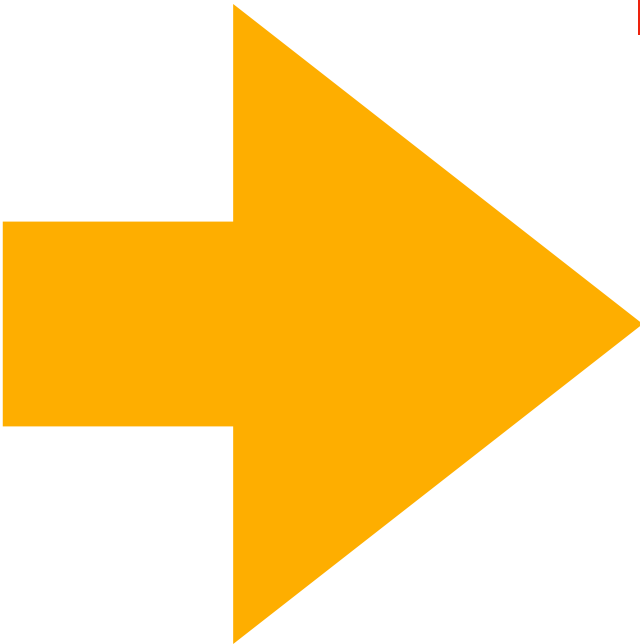
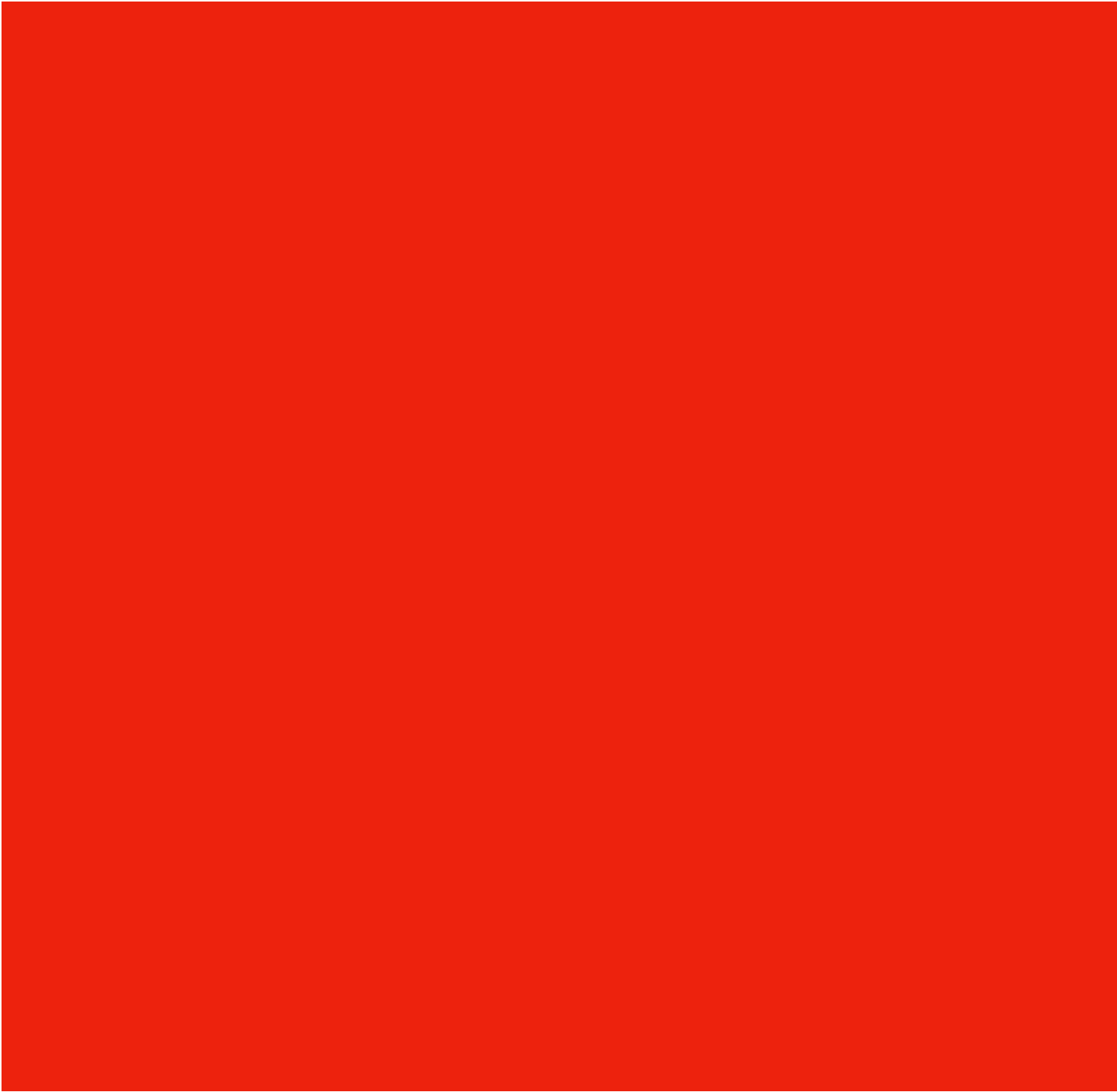


Case Studies

Stefan Wagner

Migrating to Microservices



Research Objective

Our overarching research objective is to analyze the migration process from monolithic architectures to Microservices on the basis of real-world systems in industry practice.

Research Questions

RQ1: What are intentions for migrating existing systems to Microservices?

RQ2: Which Microservices migration strategies and decomposition approaches do companies apply?

RQ3: What are the major technical and organizational challenges during a Microservices migration?

Data Collection

16 semi-structured interviews about 14 different industrial systems

Interviews were recorded and transcribed.

Data Analysis

Qualitative coding

Main categories: Intentions, Strategies, Challenges

Followed constant comparison of Grounded Theory

Companies and Participants

Company ID	Company Domain	# of Employees	Participant ID	Participant Role	Years of Experience	System ID
C1	Financial Services	1 - 25	P1	Developer	6	S1
C2	Software & IT Services	>100,000	P2	Lead Architect	30	S2
			P3	Architect	24	S3
			P4	Architect	30	S4
C3	Software & IT Services	26 - 100	P5	Architect	20	S5
			P6	Lead Developer	8	
C4	Software & IT Services	101 - 1,000	P7	Architect	9	S6
			P8	Architect	17	S7
C5	Software & IT Services	>100,000	P9	Lead Developer	7	S8
C6	Tourism & Travel	1,001 - 5,000	P10	Developer	9	S9
			P11	Data Engineer	7	
			P12	Architect	12	S10
C7	Logistics & Public Transport	101 - 1,000	P13	DevOps Engineer	5	S11
C8	Retail	5,001 - 10,000	P14	Lead Architect	9	S12
C9	Software & IT Services	101 - 1,000	P15	Architect	18	S13
C10	Retail	1,001 - 5,000	P16	Architect	22	S14

Systems and Projects

ID	Purpose	Inception	Timeframe of Migration in years	# of Services	# of People involved	Team Size	Process Model
S1	Derivatives mgmt. (banking)	Rewrite	1.75 (ongoing)	9	7	7	Scrum
S2	Freeway toll management system	Rewrite & Extension	1.5 to 2	10	10 (only devs)	5-10 (up to 40)	Individual (based on Scrum)
S3	Automotive problem management system	Rewrite & Extension	2 to 3 (ongoing)	10	50	7-9	Scrum (from Waterfall)
S4	Public transport sales system	Rewrite & Extension	2 (ongoing, exp 4)	~100	~300	6-10	Scrum, SAFe (from Waterfall)
S5	Business analytics data integration system	Greenfield	1.5 to 2 (ongoing)	6	7	7-9	Individual (based on Scrum, Kanban)
S6	Automotive configuration management system	Rewrite	0.5 (ongoing, exp 3)	60	20 (w/o cust.)	4	Scrum (from Individual)
S7	Retail online shop	Replace COTS	2.5 (ongoing)	~250	~200	6-8	Scrum, Kanban
S8	IT service monitoring platform	Cont. Evolution	2 (ongoing, exp 3)	9	15	6-10	custom
S9	Hotel search engine	Cont. Evolution	1 to 1.5 (ongoing)	~10	~50	3-6	Scrum
S10	Hotel management suite	Rewrite & Extension	0.5 to 1 (ongoing, exp 2)	20	50	1-5	Scrum, Kanban
S11	Public transport mgmt. suite	Cont. Evolution	2 to 3 (ongoing)	10	~175	5-8	Scrum
S12	Retail online shop	Replace COTS	1.5	~45	~85	6-10	Scrum, Kanban
S13	Automotive end-user service management	Rewrite & Extension	- (ongoing)	7	30	5-7	Scrum
S14	Retail online shop	Replace COTS	2.5	~175	~350	6-10	Scrum, Kanban

Intentions

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	# of Mentions
Maintainability ^a	X	X		X	X	X		X	X	X	X			X	10
Scalability ^b		X			X	X	X		X			X		X	7
Functional Requirements			X			X	X			X		X			5
Operability		X			X				X		X			X	5
Company Strategy ^c			X						X	X			X		4
Team Scalability	X		X			X	X								4
Time to Market						X			X			X	X		4
Interoperability				X		X									2
Reliability						X							X		2

Strategies

Process	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	# of Mentions
Rewrite	x	x	x	x		x	x		x	x			x		9
Strangler Pattern	x		x			x	x	x			x		x		7
Extension		x	x	x		x	x			x			x		7
Parallel Operation			x						x			x	x	x	5
Greenfield					x		x					x		x	4
COTS Replacement							x					x		x	3
Continuous Evolution								x			x				2
Decomposition															
Other (or non-systematic)	x	x	x		x	x		x	x	x	x				9
Functional Decomposition				x		x	x			x		x	x	x	7
Existing System's Structure		x	x					x	x		x				5
Domain-Driven Design				x								x		x	3

Challenges

Technical	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	# of Mentions
Decomposition		X	X	X		X		X	X	X		X			8
Lack of Expertise			X		X			X	X	X	X	X	X		8
DevOps and Automation		X			X			X	X	X	X				6
Integration of Services	X	X					X				X				4
Legacy System ^d			X			X					X		X		4
Security		X			X		X							X	4
Fault Tolerance	X				X										2
Organizational															
Mindset Change			X			X	X		X	X	X	X	X	X	9
Collaboration between Teams			X	X		X		X	X		X	X			7
Justification to Mgmt./Cust.			X		X	X					X		X	X	6
Recruiting Personnel	X		X		X			X		X				X	6
Central Governance				X					X				X		3
Volatile Requirements						X			X				X		3

Case Study – Definitions

Empirical inquiry, in which ...

- a contemporary phenomenon is investigated within its real-life context
- boundaries between phenomenon and its context are not clearly evident

Case Study – Definitions

Empirical inquiry, which ...

- copes with the technical distinctive situation in which there will be many more variables of interest than data points,
- relies on multiple sources of evidence (data needs to be converged) and
- benefits from prior development of theoretical propositions to guide data collection and analysis (or generates a new theory).

Experiment

Smaller sample
Less control
Realistic environment

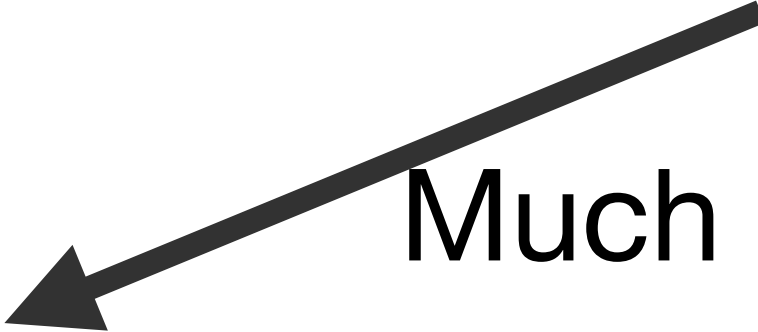
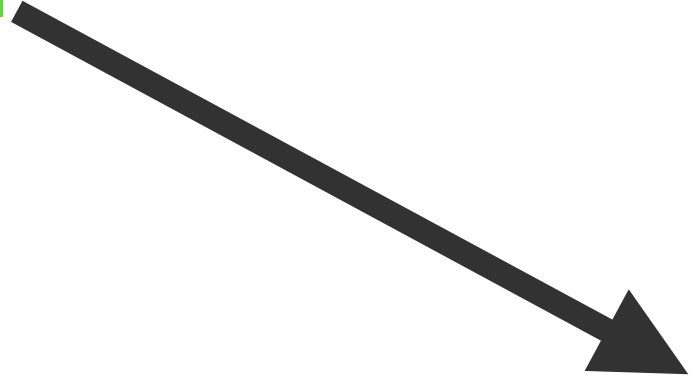
Survey

Much smaller sample
Wider range of data collection
Potentially direct observation

Case Study

Less focus on change

**Action
Research**



Types of Case Studies

Improving

- The studied phenomenon is improved in some way
- Close to action research
- Example: Does the introduction of user stories lead to less effort in requirements engineering?
- Example: Does this new hazard analysis technique find more hazardous scenarios?

Exploratory

- Criteria or parameters instead of purpose
- Example: Are static analysis tools used and if yes, why?
- Example: What do CMM level 3 organisations have in common?

Types of Case Studies

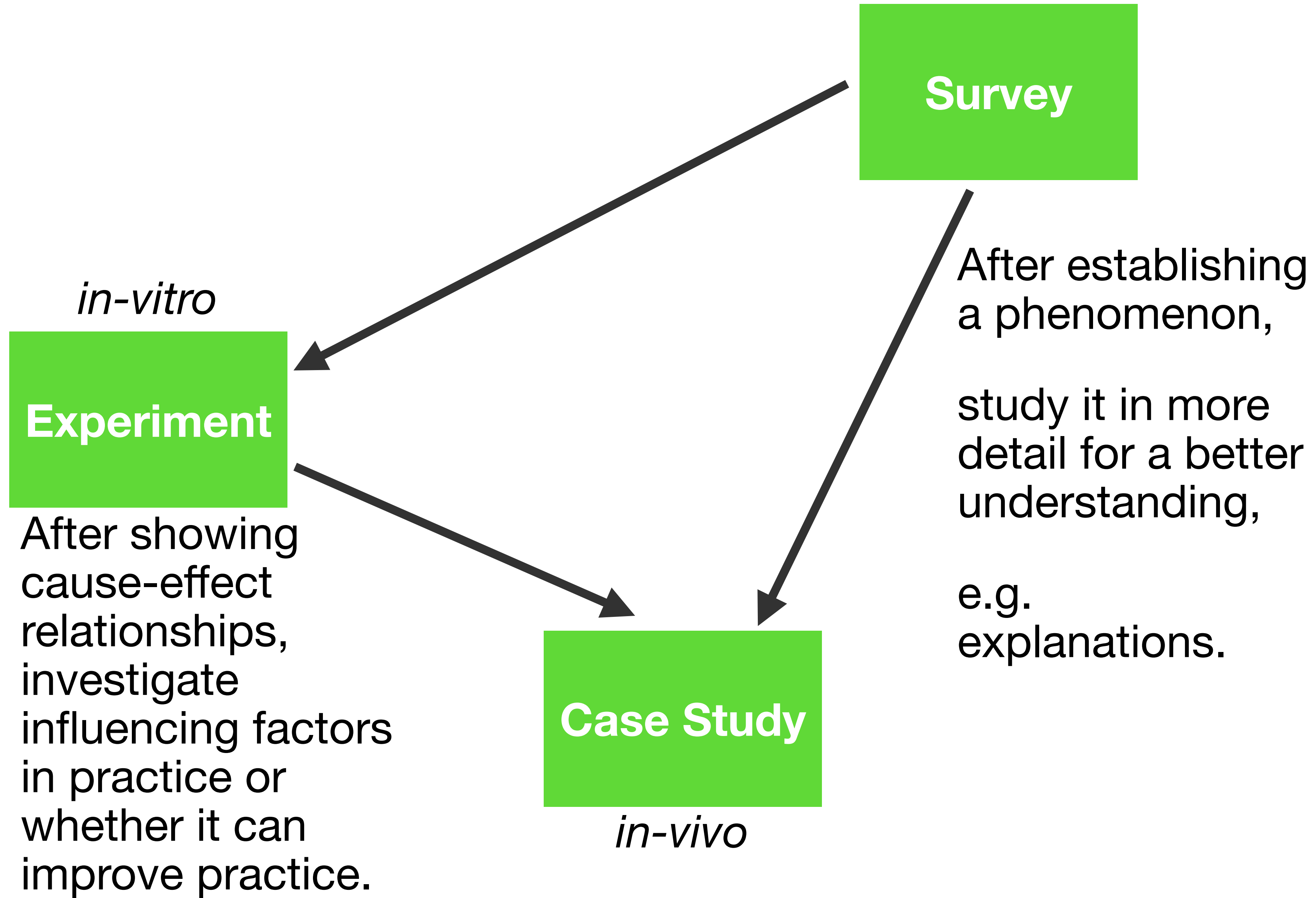
- Adjudicates between competing explanations
- Example: Does the usage of static analysis tools reduce the number of field defects?
- Rival theories: existing architectures are useful for anchoring, vs. existing architectures are over-constraining during RE

Descriptive

- Describes sequence of events and underlying mechanisms
- Example: How does pair programming actually work?
- Example: How is static analysis used in practice?

Overview of Research Strategy Characteristics

	Experiment	Survey	Case Study	Action Research
Primary Objective	Explanatory	Descriptive	Exploratory	Improving
Primary Data	Quantitative	Quantitative	Qualitative	Qualitative
Design Type	Fixed	Fixed	Flexible	Flexible



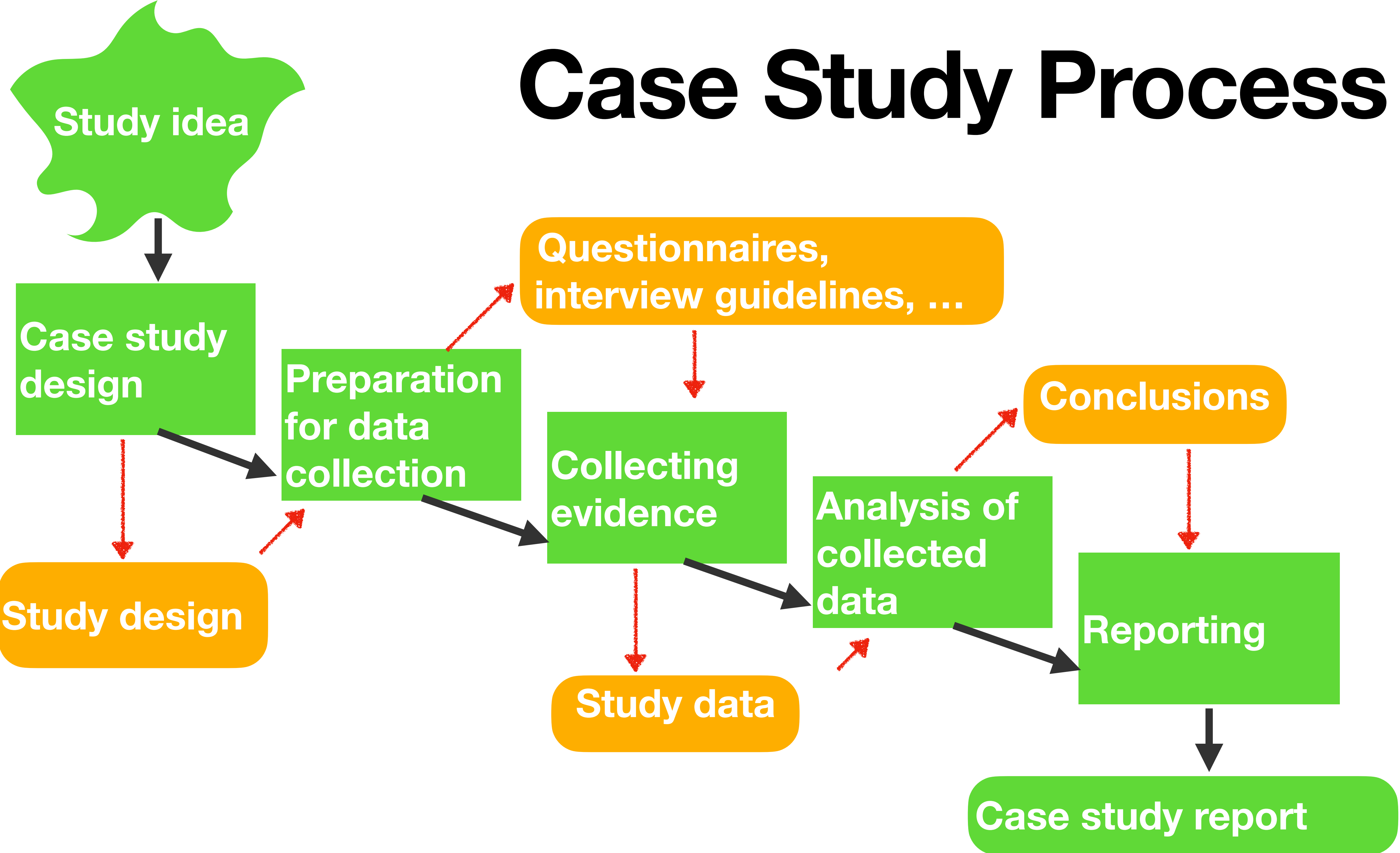
Example

Let's assume we have found our new static analysis tool in our experiments to be quite effective and efficient in our experiments.

Now we want to understand static analysis and especially our new tool in a realistic setting.

Therefore, we need to perform a case study that helps us to understand what are current problems with static analysis and how well developers could work with our tool.

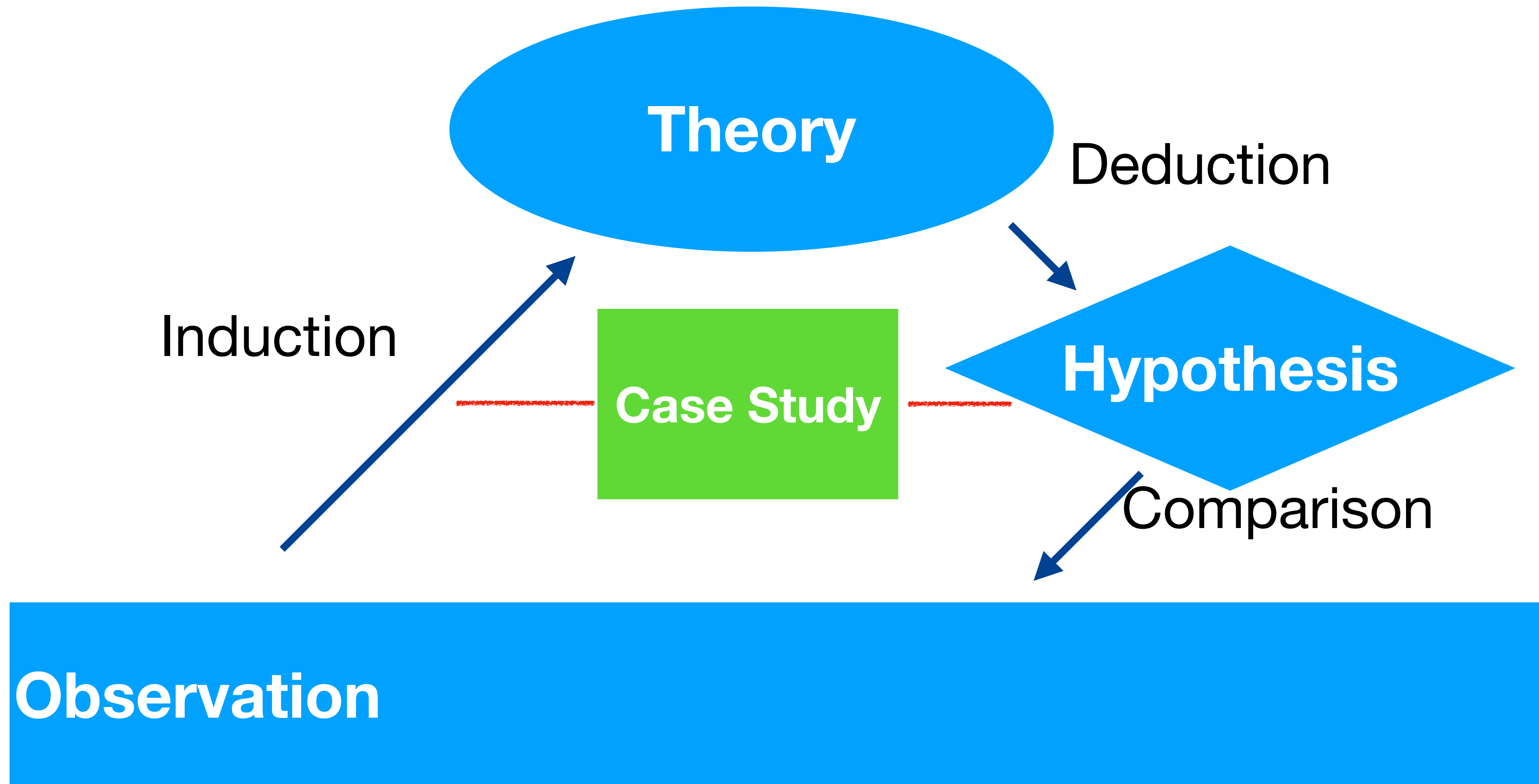
Case Study Process



Rationale and Objective of the Study

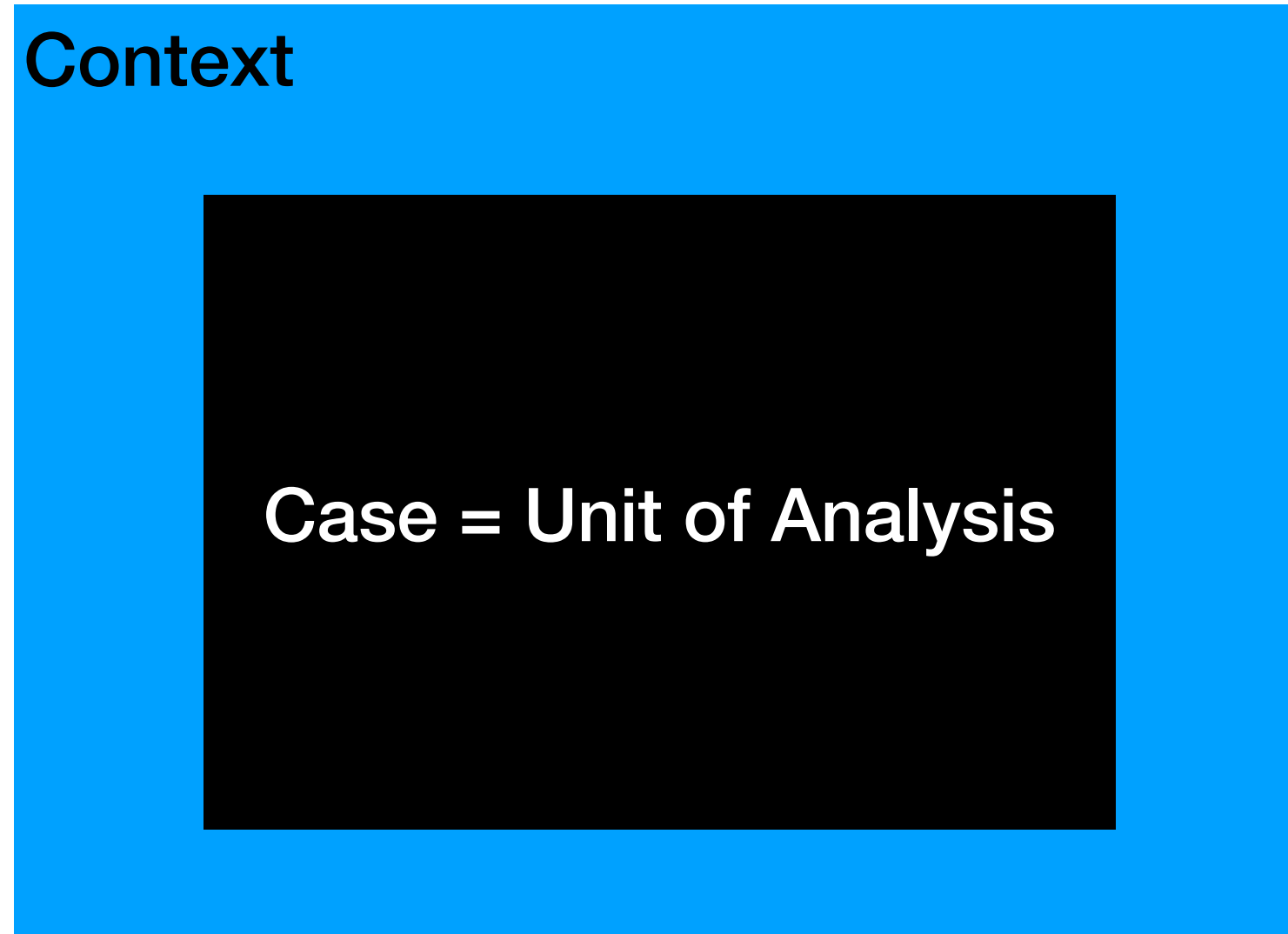
Rationale: A research gap or practical problem

Objective: What do we want to achieve with the study?

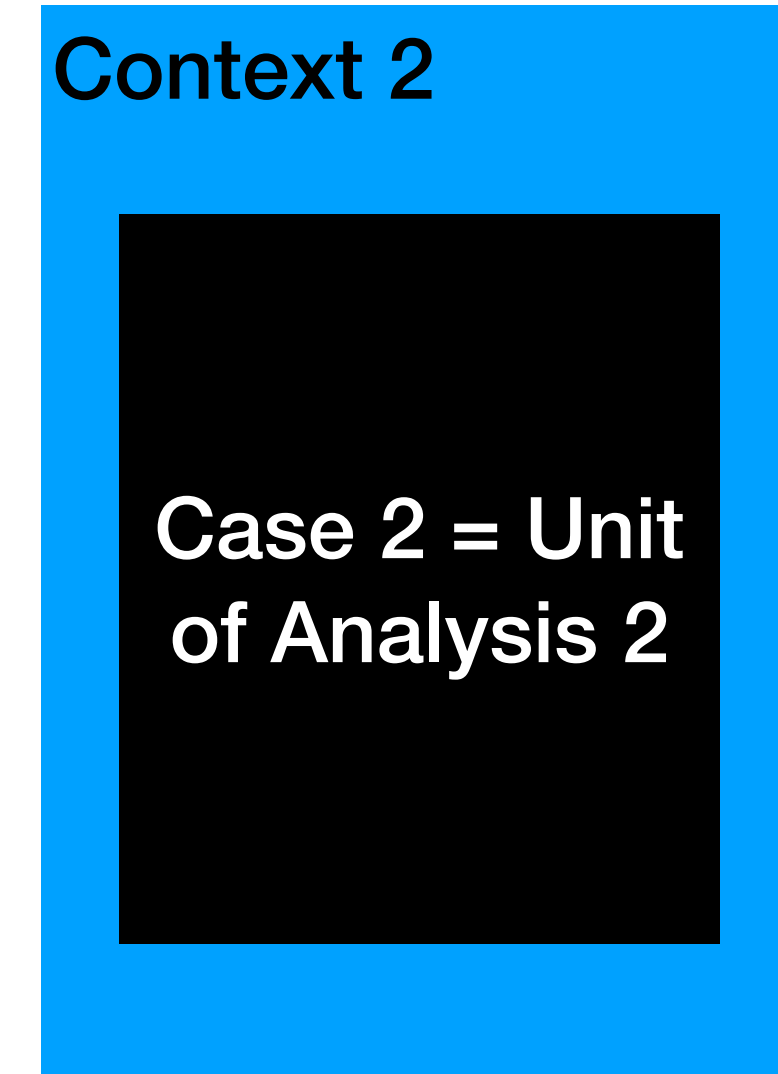
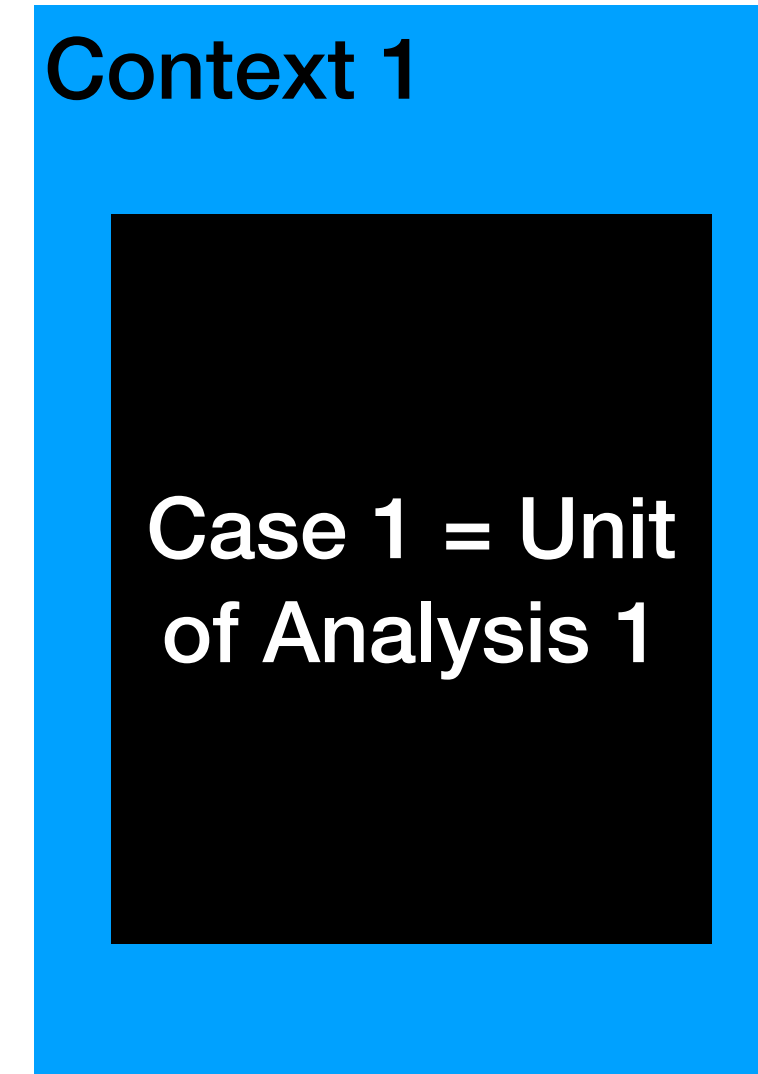


Holistic case study

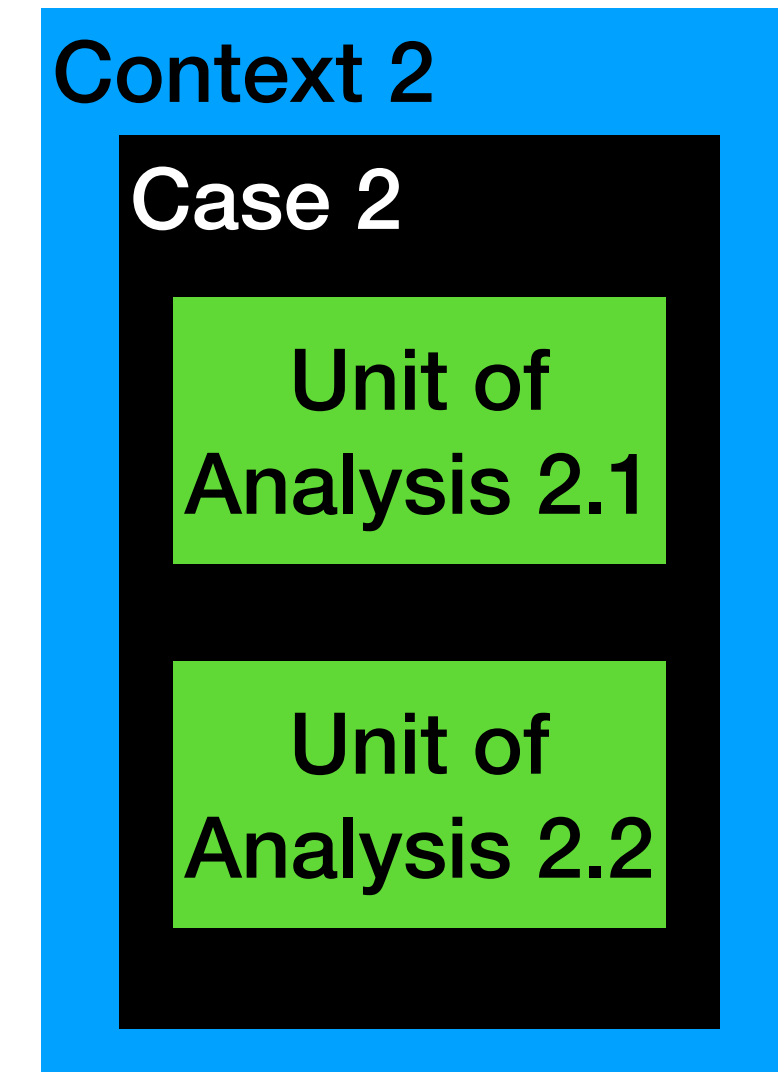
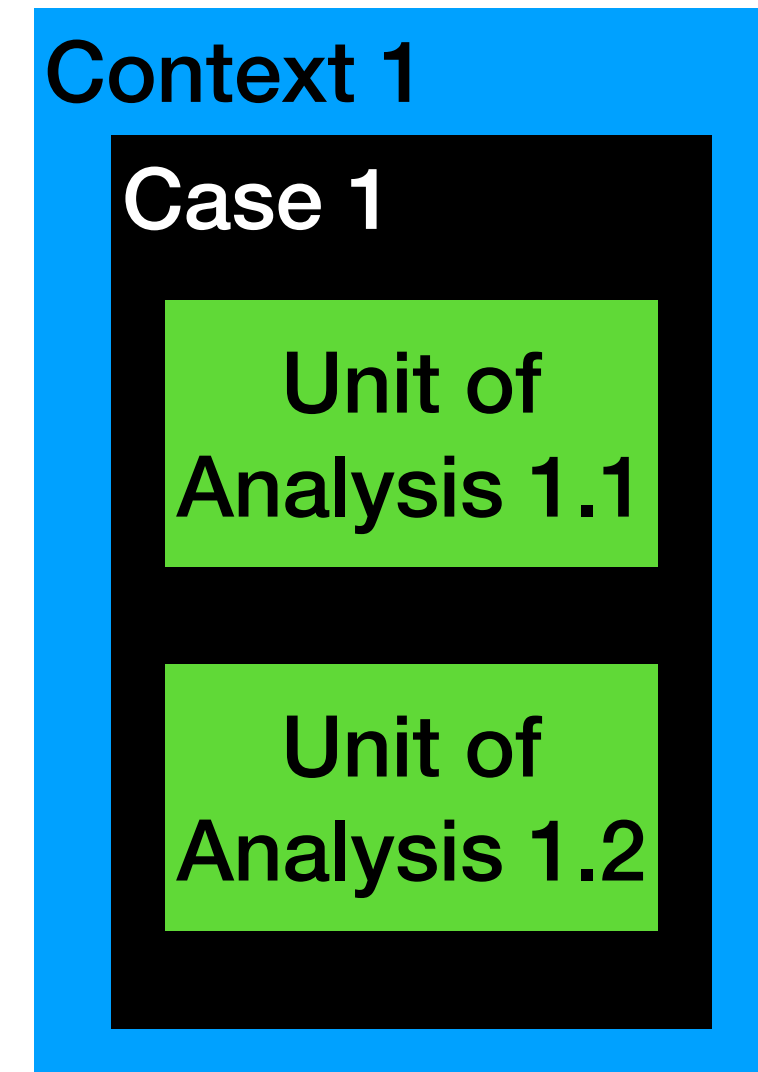
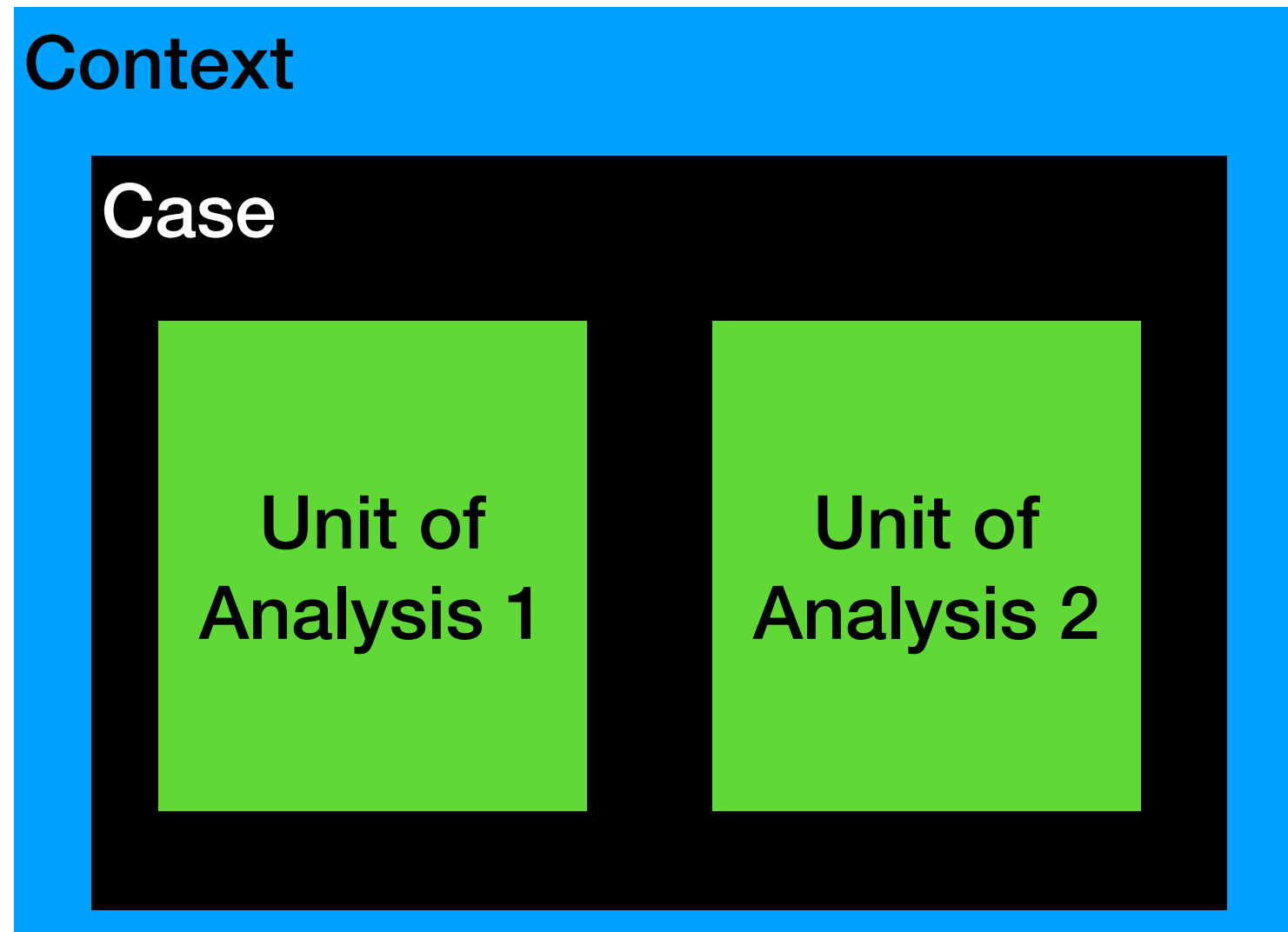
Single-case study



Multiple-case study



Embedded case study



Theoretical Framework

What is the theoretical frame of reference?

Literature study

Cognitive theories

Social theories

Organisational theories

...

Research Questions and Hypotheses

Every case study should have research questions.

But not every case study can have hypotheses.

Case Selection

Extreme/deviant: To obtain information on unusual cases, which can be especially problematic or especially good in a more closely defined sense.

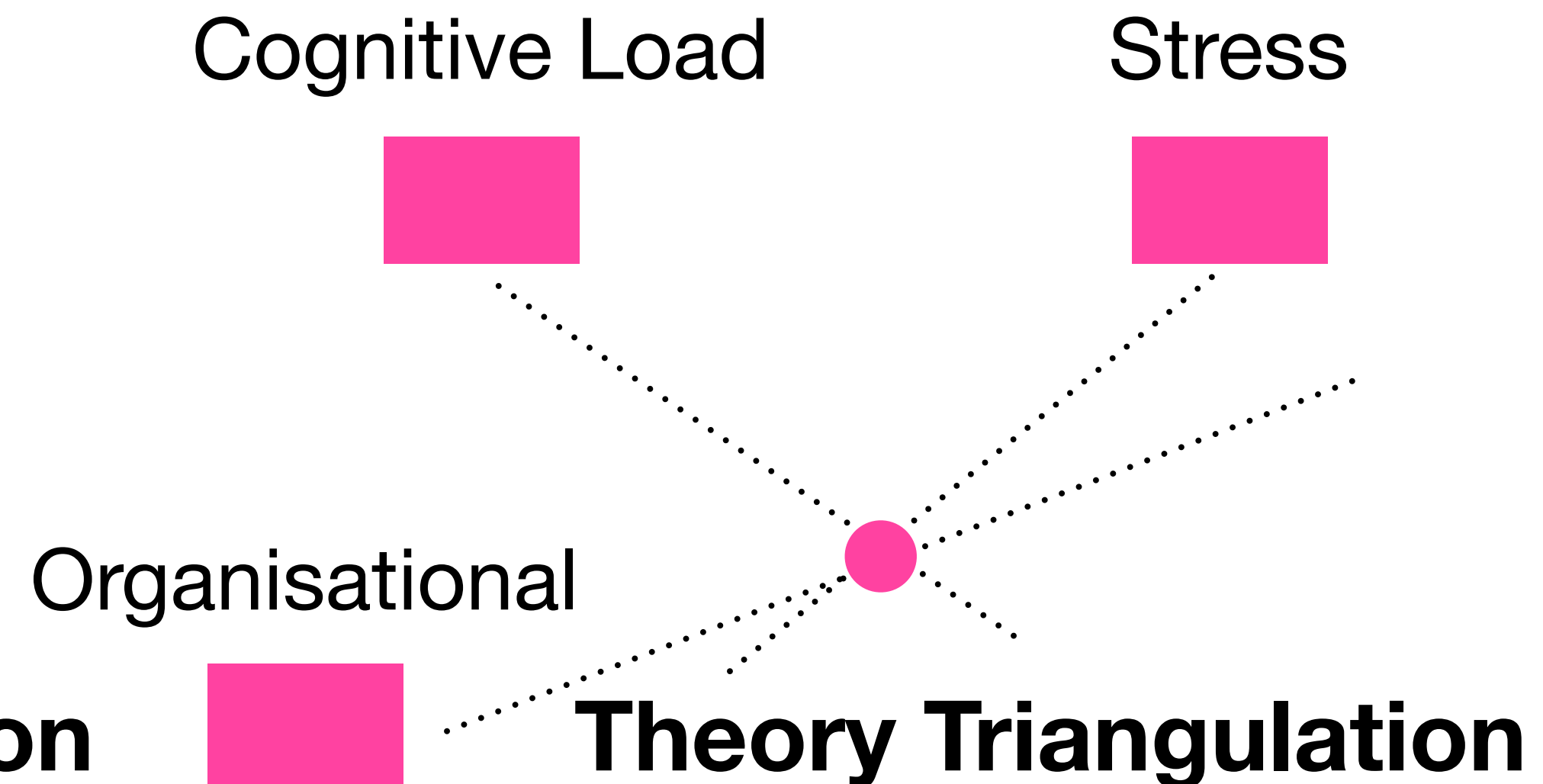
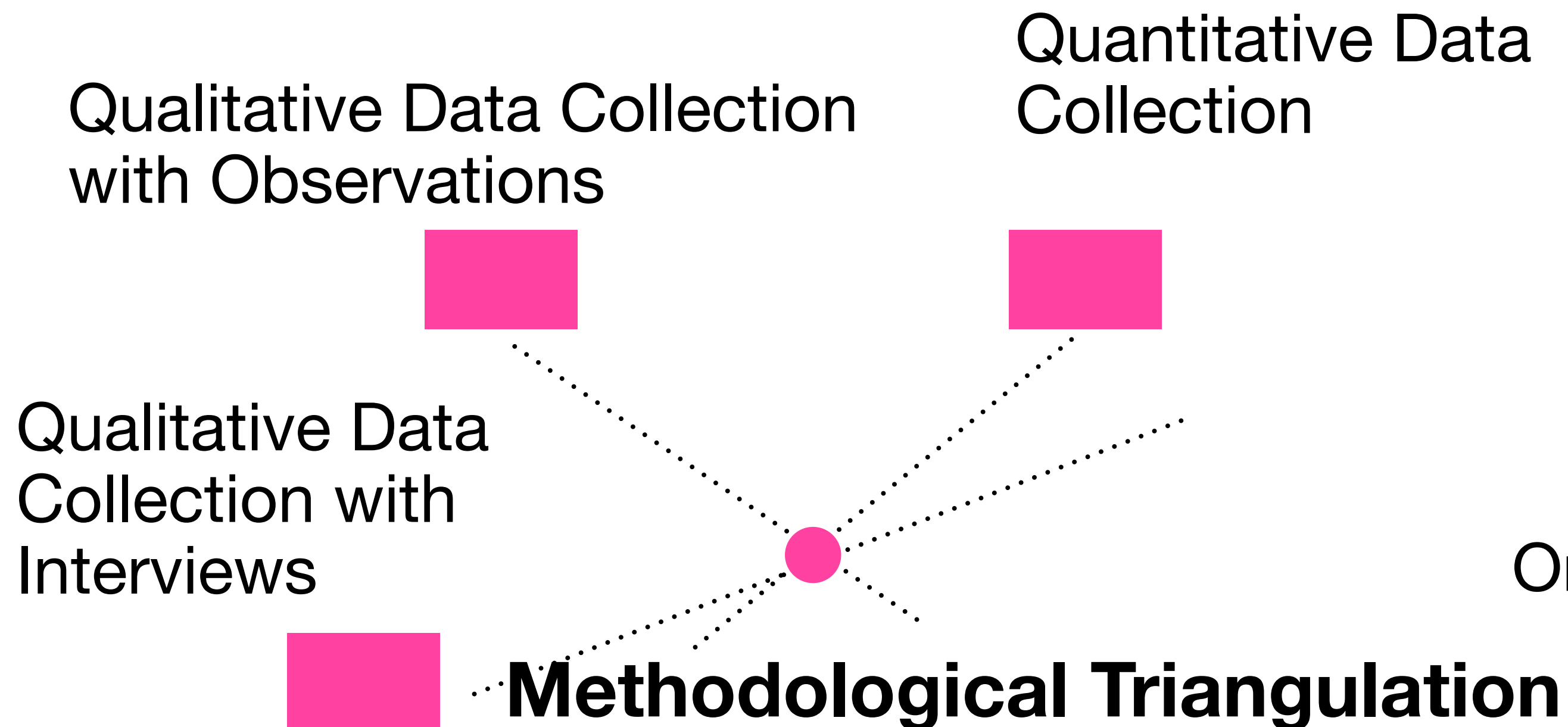
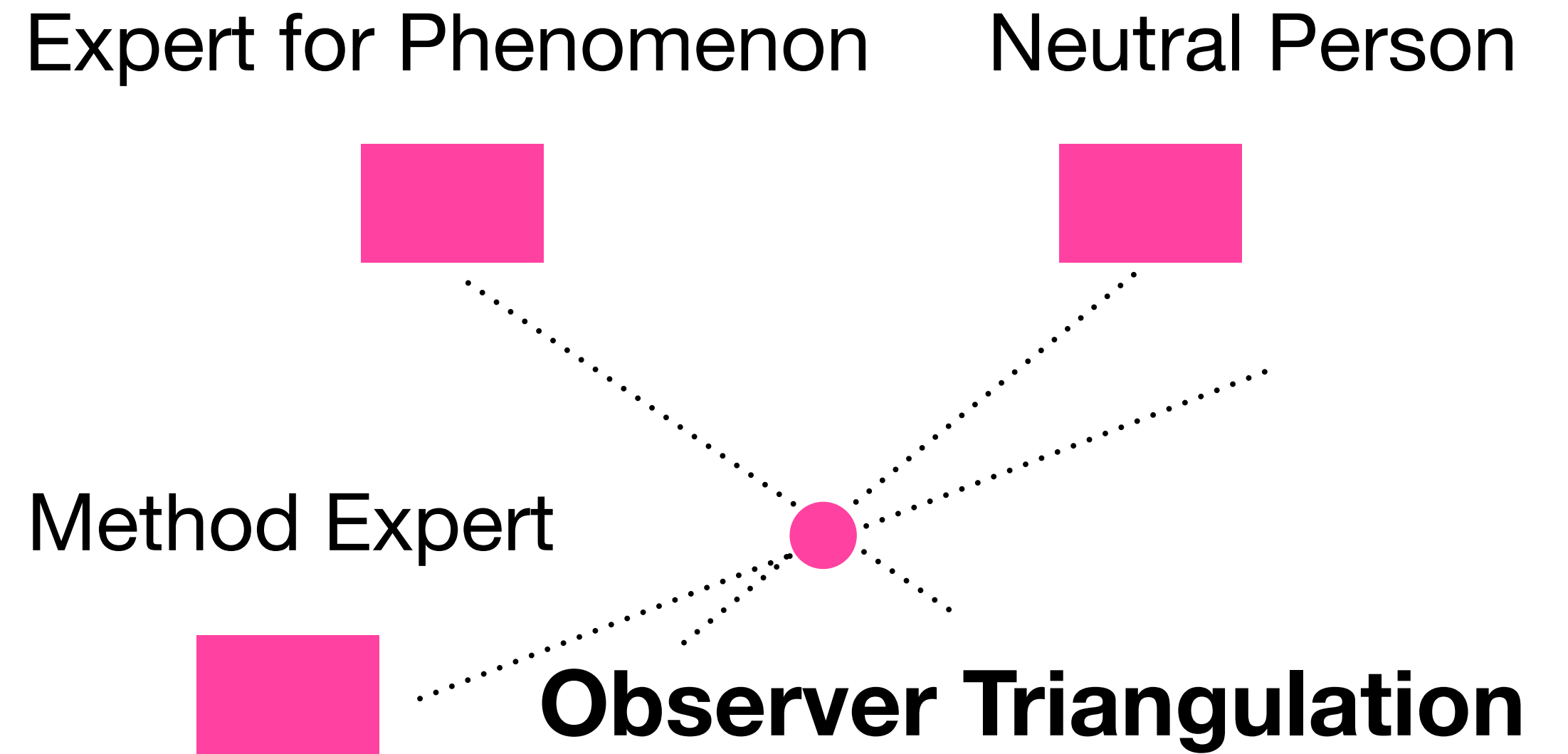
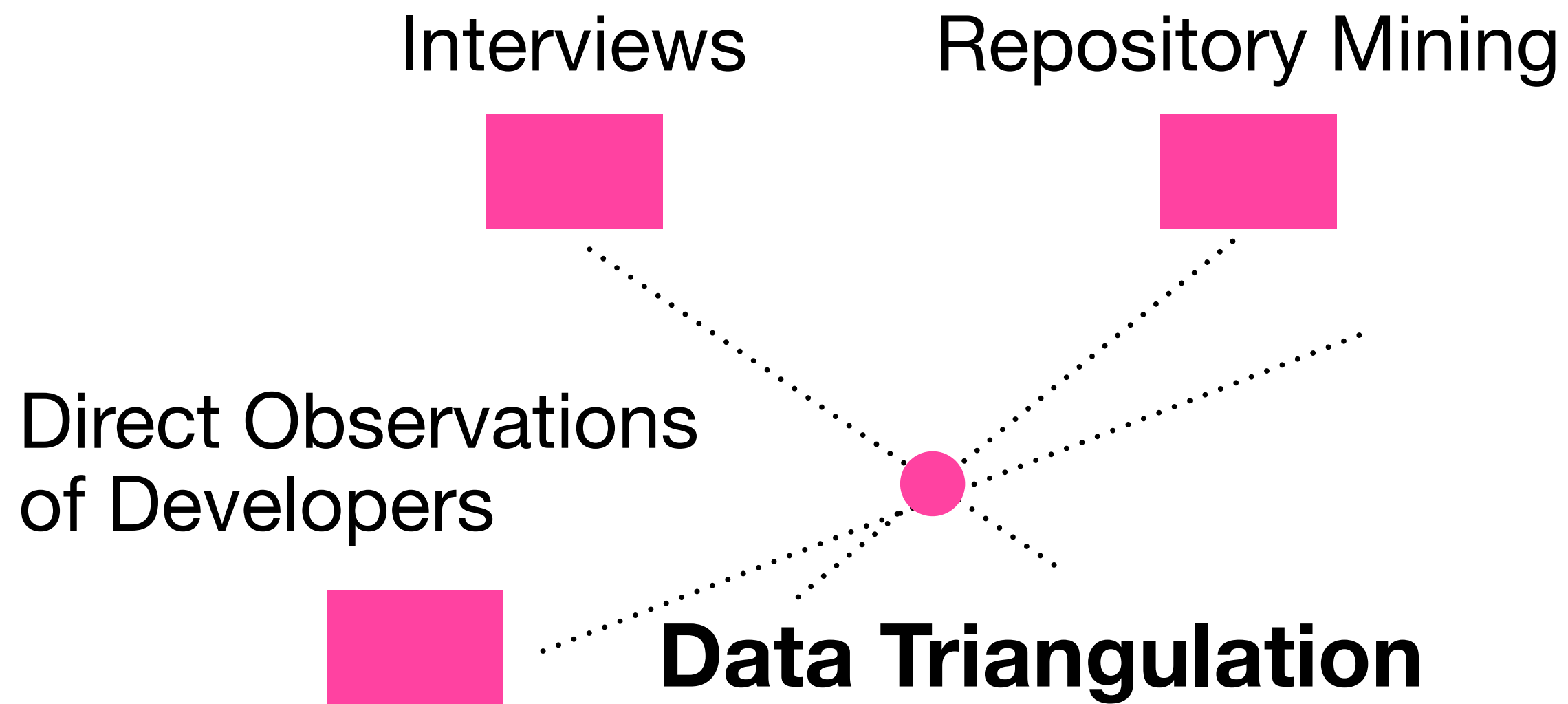
Maximum variation: To obtain information about the significance of various circumstances for case process and outcome (e.g., three to four cases that are very different on one dimension: size, form of organisation, location, budget).

Critical: To achieve information that permits logical deductions of the type, “If this is (not) valid for this case, then it applies to all (no) cases.”

Paradigmatic: To develop a metaphor or establish a school for the domain that the case concerns.

From: Flyvbjerg (2007) as cited in Runeson et al. (2012)

Triangulation



Threats to Validity

Threat to	Case study tactic	Phase of research in which tactic occurs
Construct Validity	<ul style="list-style-type: none"> • Use multiple sources of evidence • Establish chain of evidence • Have key informants review draft report 	<p>Data collection</p> <p>Data collection Composition</p>
Internal Validity	<ul style="list-style-type: none"> • Do pattern-matching • Do explanation-building • Address rival explanations • Use logic models 	<p>Data analysis</p> <p>Data analysis</p> <p>Data analysis</p> <p>Data analysis</p>
External Validity	<ul style="list-style-type: none"> • Use theory in single-case studies • Use replication logic in multiple-case studies 	<p>Research design</p> <p>Research design</p>
Reliability	<ul style="list-style-type: none"> • Use case study protocol • Develop case study database 	<p>Data collection</p> <p>Data collection</p>

Design Checklist

1. What is the case and its units of analysis?
2. Are clear objectives, preliminary research questions, hypotheses (if any) defined in advance?
3. Is the theoretical basis – relation to existing literature or other cases – defined?
4. Are the authors' intentions with the research made clear?
5. Is the case adequately defined (size, domain, process, subjects...)?
6. Is a cause-effect relation under study? If yes, is it possible to distinguish the cause from other factors using the proposed design?
7. Does the design involve data from multiple sources (data triangulation), using multiple methods (method triangulation)?
8. Is there a rationale behind the selection of subjects, roles, artefacts, viewpoints, etc.?
9. Is the specified case relevant to validly address the research questions (construct validity)?
10. Is the integrity of individuals/organisations taken into account?

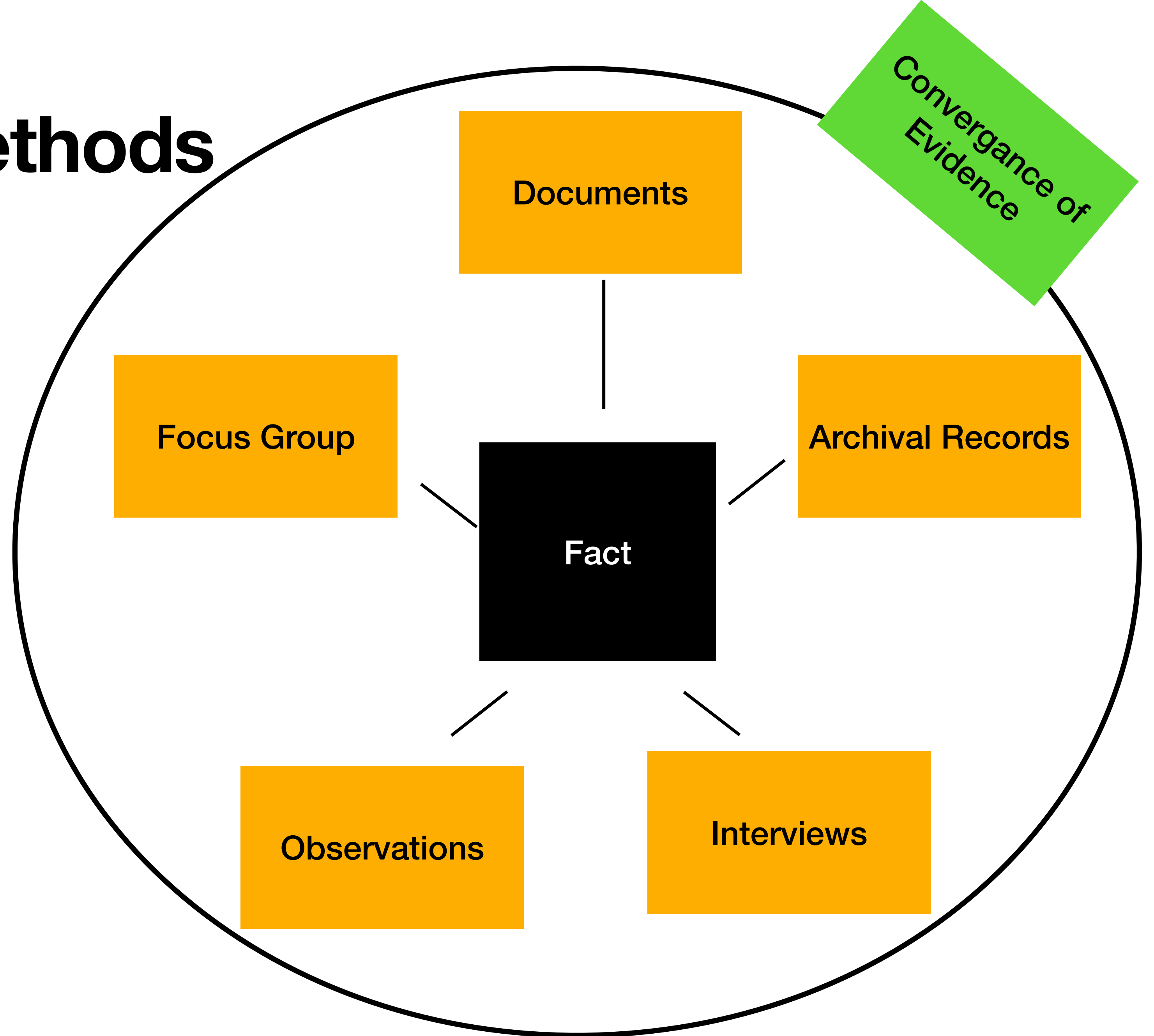
Refactoring

Let's assume that our experiment on the effects of specific refactorings did not show any statistically significant results. Yet, there are many practitioners advocating it. Hence, we want to understand its usage and effects in practice!

Make a rough case study design!

Data Collection Methods

- qualitative, quantitative, combination
- triangulation “to bring the data together”



Principles of Data Collection

- Use multiple data sources: Triangulation, i.e. searching convergent findings from different sources (→ Increase construct validity)
- Create a case study database
 - Content
 - Case study notes (clear & available for later use)
 - Case study documents
 - Tabular materials (collected & created)
 - Narratives (initial open-ended answers to the study questions suggested by investigators)
 - Separate from the final report to be written

Principles of Data Collection

- Maintain a chain of information: Explicit documentation of the traceability between research questions and case study procedures.
 - Ensure the collection of all data required for answering the research questions.
 - Justify the collection of each data item.
- Design and use the case study protocol for supporting data collection and analysis.
- Storage of actual data in the data base for later reviews including elicitation circumstances.
- Explicit citation of data sources and data base location in the final report and conclusions

Classification of Data Sources

First degree

Interviews

Focus groups

Observations with
„think aloud“

Second degree

Screen capturing

Third degree

Analysis of work
artefacts

Repository
mining

Interviews

	Unstructured	Semistructured	Fully Structured
Typical Foci	How individuals qualitatively experience the phenomenon	How individuals qualitatively and quantitatively experience	Researcher seeks to find relations between constructs
Interview Questions	Interview guide with areas to focus on	Mix of open and closed questions	Closed questions
Objective	Exploratory	Descriptive and explanatory	Descriptive and explanatory

Postinterview Activities

- Consolidate notes and/or transcribe recording
- Give notes/transcription to interviewee for feedback

Observations

Screen capturing

Observations
with „think
aloud“

Video recordings

Archival Data

Meeting minutes

Requirements specifications

Test specifications

Source code

Project plan

Organisational
charts

Financial records

Reports

Metrics

Sales figures

Defect counts

Time sheets

Integration times
and failures

Project costs

Data Collection Checklist

1. Is a case study protocol for data collection and analysis derived (what, why, how, when)? Are procedures for its update defined?
2. Are multiple data sources and collection methods planned (triangulation)?
3. Are measurement instruments and procedures well defined (measurement definitions, interview questions)?
4. Are the planned methods and measurements sufficient to fulfil the objective of the study?
5. Is the study design approved by a review board, and has informed consent obtained from individuals and organisations?
6. Is data collected according to the case study protocol?
7. Is the observed phenomenon correctly implemented (e.g. to what extent is a design method under study actually used)?
8. Is data recorded to enable further analysis?
9. Are sensitive results identified (for individuals, the organisation or the project)?
10. Are the data collection procedures well traceable?
11. Does the collected data provide ability to address the research question?

Refactoring

How could we apply these different data collection techniques in our refactoring case study?

Develop an interview guide that you would use for interviewing developers in a company!

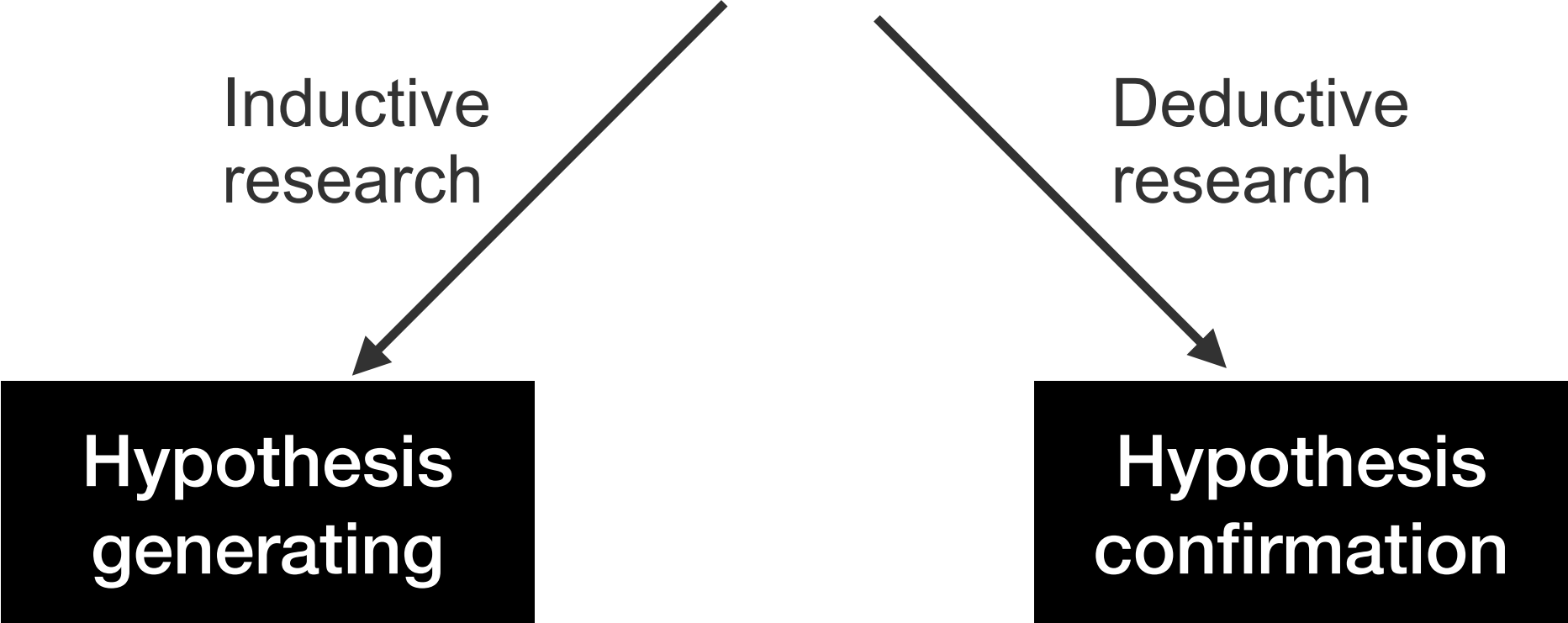
Data analysis

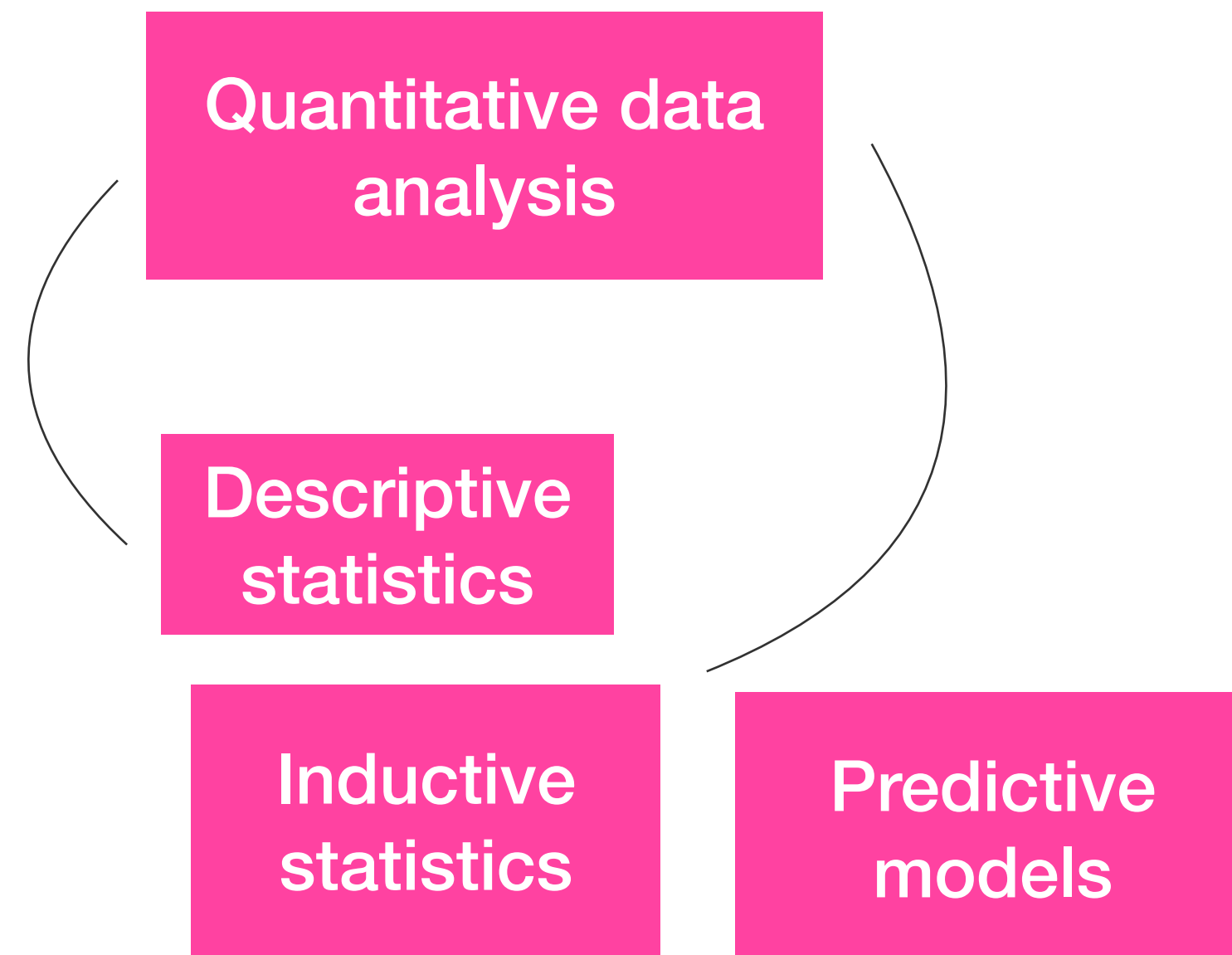
Inductive
research

Deductive
research

**Hypothesis
generating**

**Hypothesis
confirmation**





Qualitative data analysis

**Grounded
Theory**

**Qualitative
content
analysis**

...

**Discourse
analysis**

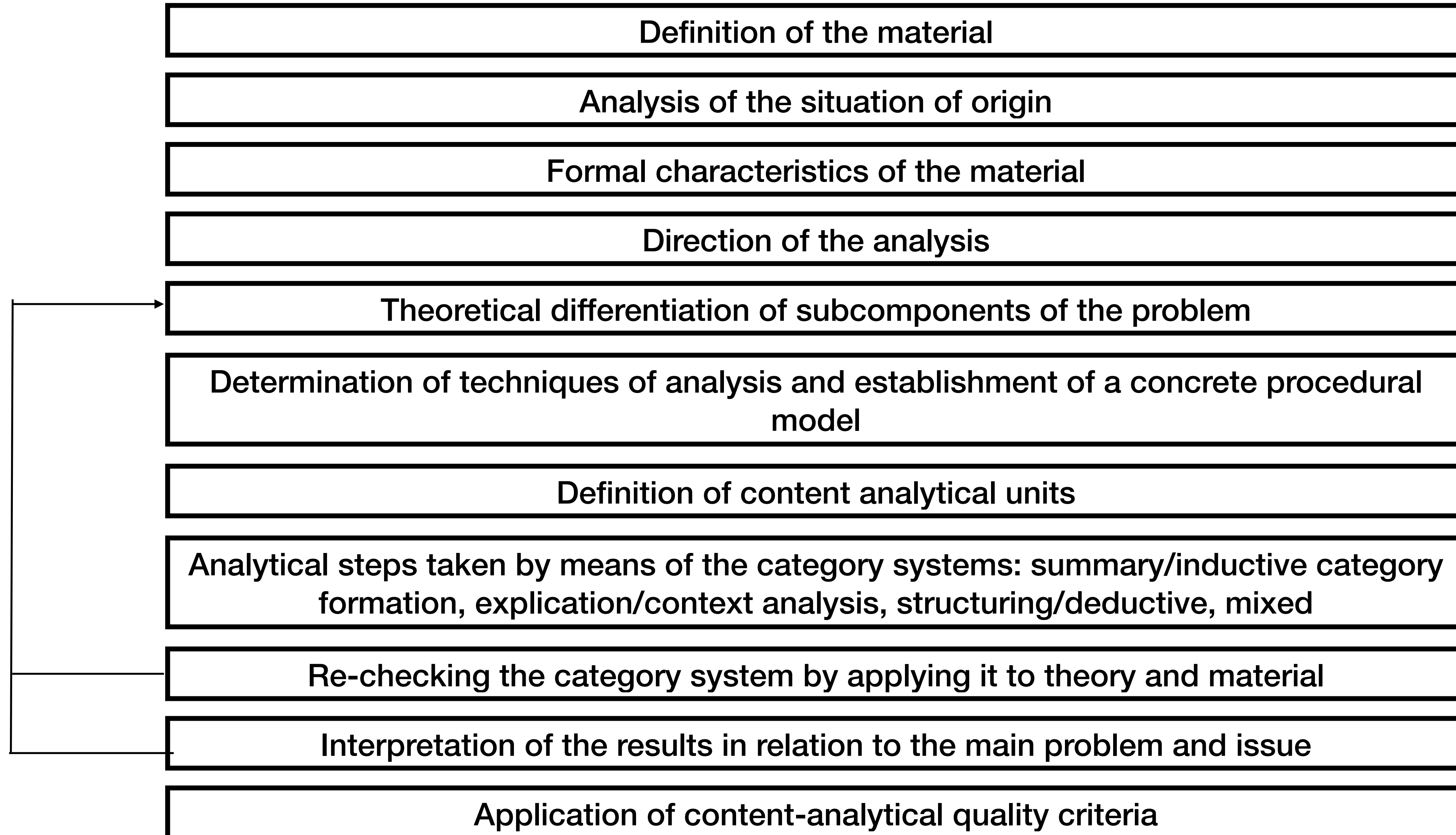
Hermeneutics

Ethnography

Qualitative Content Analysis

- Proposed by Mayring
- Its aim is to analyze protocols of communication
- It can be applied in various disciplines, such as psychology, sociology, and linguistics.
- Depending on the discipline, the unit of analysis may be quite different. In SE, we usually start at the phrase or sentence level.

Qualitative Content Analysis



What exactly is part of the text to analyse?

Who was the author? Which background?

How did we get to the text (e.g. transcription)?

On what level should the output of the analysis be?

Can we theoretically structure the problem?

In which concrete steps do we plan to perform the analysis?

Do we look at single words, phrases, sentences, pages?

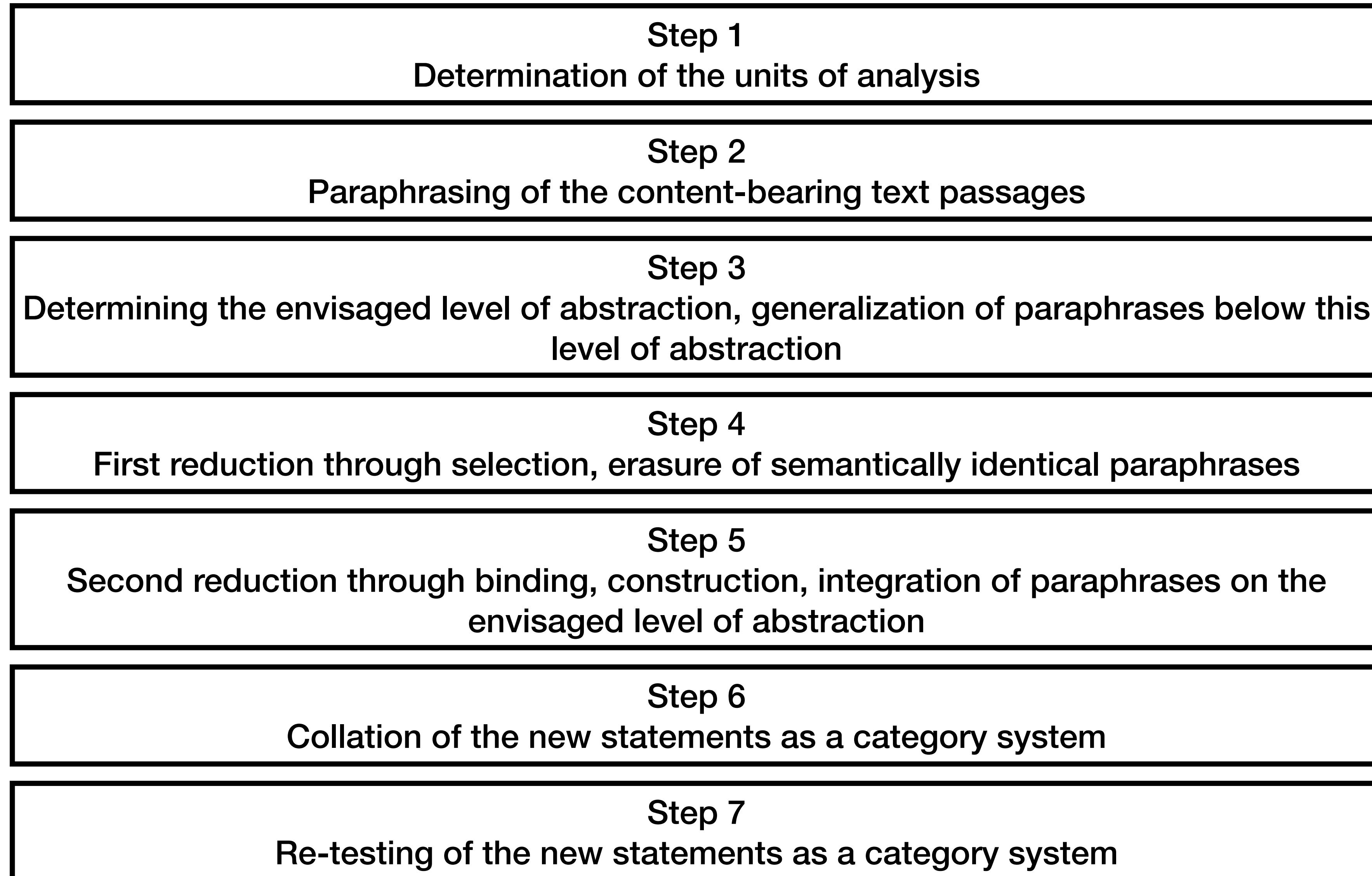
Does the category system fit to the theory and the material?

Are there differences between different coders?

Basic Forms of Interpretation

- **Summary:** The objective of the analysis is to reduce the material in such a way that the essential contents remain, to create through abstraction a comprehensive overview of the base material which is nevertheless still an image of it.
- **Explication:** The objective of the analysis is to provide additional material on individual doubtful text components (terms, sentences...) with a view to increasing understanding, explaining, interpreting the particular passage of text.
- **Structuring:** The objective of the analysis is to filter out particular aspects of the material, to give a cross-section through the material according to pre-determined ordering criteria, or to assess the material according to certain criteria.

Summarizing



From: Mayring (2014)

Summarizing

Case	Page	Paraphrase	Generalisation	Reduction
A	43	P1: Dev. was shocked by the many warnings presented to her after running the tool.	Overwhelmed by number of warnings	C1: Developer is overwhelmed by the number, diversity and summaries from the static analysis report.
A	44	P2: Dev. did not know where to start understanding the warnings from FindBugs.	Overwhelmed by diversity of warnings	
A	44	P3: Dev. could not make sense of lengthy summaries	Overwhelmed by the summaries of warnings	

Context Analysis

Step 1

Determination of evaluation unit, i.e., establishing the portion of text to be interpreted

Step 2

Lexical-grammatical definition of the portion of text involved

Step 3

Determining the additional explication material permissible

Step 4

Collation of the material

Narrow context analysis: direct text environment

Broad context analysis: additional material beyond the limits of the text

Step 5

Phrasing of interpretative paraphrase(s)

Step 6

Testing the sufficiency of the explication

Deductive Category Assignment

Step 1

Research question, theoretical background

Step 2

Definition of the category system (main categories and subcategories) from theory

Step 3

Definition of the coding guideline (definitions, anchor examples and coding rules)

Step 4

Material run-through, preliminary codings, adding anchor examples and coding rules

Step 5

Revision of the categories and coding guideline after 10–50 % of the material

Step 6

Final working through the material

Step 7

Analysis, category frequencies and contingencies interpretation

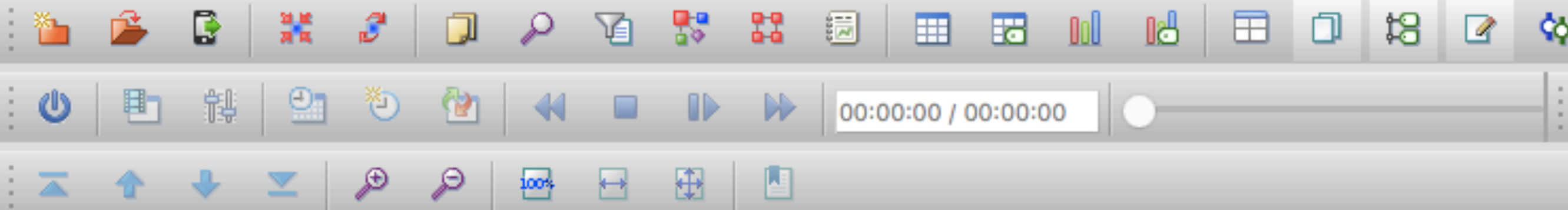
Deductive Category Assignment

Variable	Value	Definition	Anchor sample	Coding rules
Static analysis usage	U1: high level of usage	Frequent daily use of several static analysis tools	„I check every change with FindBugs before I commit.“	Both aspects must be met.
	U2: medium level of usage	Use of some tool at least every week	„I apply Coverity before I present my work to the PO.“	
	U3: low level of usage	Use of some tool from time to time	„We don't check our code regularly but before we release.“	
	U4: usage not inferable	No clear statement from the developer	„We do have some static analysis tools.“	

Inter-Coder Agreement

- Give text to be analysed to second person and compare outcome.
- A second coder gets all the category definitions and rules for a recoding.
- Light-weight: Review of the coding of the first coder

We can also calculate the inter-coder agreement, e.g. with Cohen's kappa.



Liste der Dokumente

Dokumente	182
Interviews	157
B01 Jan	36
B02 Maria	27
B03 Sarah	28
B04 Hans	35
B05 Lukas	31
Bilder Klimawandel	0
Überschwemmung3	0
Überschwemmung2	0
Überschwemmung1	0
Webseiten	2
Klimakulturen	2
www.klimabewusstsein.de	0
www.ipcc.ch	0

Liste der Codes

Codesystem	182
WP - Größte Weltprobleme	8
Klima	6
Ressourcenknappheit, -verteilung, Armut	4
Egoismus, fehlende Gemeinsamkeit	2
religiöse, kulturelle Konflikte	2
Krieg	2
Globalisierung	1
Schnellebigkeit	1
EI - Gesellschaftliche Einflussnahme	44
KK - Einfluss des Konsums auf Klimawandel	8
UD - Ursachen für die Diskrepanz	11
REL - Pers. Relationen zur globalen Entwicklung	4
PH - Pers. Handeln	26
VER - Pers. Verantwortungsübernahme	5
LER - Erlernbarkeit	29

Dokument-Browser: B01 Jan

1 Interview mit Jan

2 I: Ok. Ja hallo.

3 B: Ja hallo. schönen guten Tag.

4 I: Schön, dass du Zeit für mich hast. Wir haben ja schon telefoniert und du weißt ja schon worum es geht.

5 B: Ja ich weiß so, du machst ein Interview. Und du willst mir ein paar Fragen stellen, wo ich was zu sagen soll.

6 I: zum Umweltverhalten. Ok, dann fang ich einfach mal mit der ersten Frage an. Und die lautet, was sind aus deiner Sicht die größten Probleme der Welt im 21. Jahrhundert generell?

7 B: Joa, das sind die jetzt die aktuellen Probleme die wir haben. Es gibt ja da die materiellen Probleme, dass wir halt Energieprobleme haben, wie regeln wir unsere Energieversorgung für die Zukunft? Wo gehen wir hin? Irgendwann sind die fossilen Sachen sind alle. Und welche Alternativen sprich diese Bioenergie und Solarenergie ähm...weniger Vor- und Nachteile. Wasserstoff ist noch nicht soweit, dass man sagen könnte dass es ausgereift ist. Das gibt es ja noch nicht aufm Markt. Das sehe ich als ein ganz großes Problem. Dann natürlich auch Klimaveränderungen, wobei da auch viel Hysterie drin ist, wir brauchen jetzt innerhalb von einem Jahr nicht alles rumdrehen, was in den letzten 60 Jahren in die Atmosphäre geblasen worden ist, also das sollte man nicht so hektisch betrachten. Oder dann ist das natürlich so, dass wir auch ein emotionales Problem haben und zwar ist die Zeit zu schnelllebig geworden. Wir nehmen uns zu wenig Zeit für wichtige Sachen wie Gespräche für Familie und diverser und die moderne Elektronik die es so gibt, zum Beispiel Handys, klingeln rund um die Uhr und wir kommen gar nicht mehr zur Ruhe. Das ist auch ein Problem, wo vielleicht auch die eine oder andere Krankheit irgendwo entsteht.

8 I: Gut, sonst noch irgendwelche Probleme was jetzt nicht mit dem Klimawandel zu tun hat, generelle Probleme, große Probleme die es auf der Welt gibt.

9 B: Ja gut, wir haben noch Hungersnöte, die Kriege, das sind auch große Probleme, die sind weiter weg

..Ressourcenknappheit, -verte

WP - Größte Weltprobleme

..Klima

..Schnellebigkeit

Analysis Checklist

1. Is the analysis methodology defined, including roles and review procedures?
2. Is a chain of evidence shown with traceable inferences from data to research questions and existing theory?
3. Are alternative perspectives and explanations used in the analysis?
4. Is a cause-effect relation under study? If yes, is it possible to distinguish the cause from other factors in the analysis?
5. Are there clear conclusions from the analysis, including recommendations for practice/ further research?
6. Are threats to the validity analyzed in a systematic way and countermeasures taken? (Construct, internal, external, reliability) question?

Refactoring

We have a transcript of one interview with one developer regarding her use of refactorings in her last project. Try to do a summarizing qualitative analysis on it.

Reporting the Context

Facet	Element
Product	Maturity Quality Size System type Customization Programming language
Processes	Activities Workflow Artifacts
Practices, tools, technique	CASE tools Practices and techniques
People	Roles Experience
Organization	Model of overall organization Organizational unit Certification Distribution
Market	Number of customers Market segments Strategy Constraints

Reporting Checklist

1. Are the case and its units of analysis adequately presented?
2. Are the objective, the research questions and corresponding answers reported?
3. Is the report suitable for its audience, easy to read and well structured?

Desired Skills of the Investigators

- Mastering the art of asking insightful questions and skillfully interpreting the answers, empowering you with a deeper understanding of the subject matter.
- Being a good listener without being trapped by preconceptions or ideologies.
- Being adaptive and flexible, perceiving newly encountered situations as opportunities rather than threats, and being open to change.
- Having a good grasp of the issue being studied and not missing essential clues.
- Understanding when a deviation is acceptable.
- Not mechanically recording data, but interpreting information in real-time and being able to react to contradictions among sources of evidence.
- Avoiding bias and not using a case study to substantiate a preconceived position.
- Being open to contrary findings.

Criticism

Please remember the following points:

- **Criticism:** Lack of systematic handling of data (Lack of rigor!)
- **Response:** It is crucial to address this concern by systematically reporting all data and procedures involved in the research.

- **Criticism:** Little basis for scientific generalization!
- **Response:** The purpose of the research is to generalize to theoretical propositions, and not to a population as in statistical research. This involves analytic generalization rather than statistical generalization.

- **Criticism:** Takes too long, resulting in massive, unreadable documents.
- **Response:** The time taken for the research process depends on the specific research question at hand. The analysis and documentation will be carried out based on the choices made by the investigators.

References

- B. Flyvbjerg. Five misunderstandings about case-study research. In *Qualitative Research Practice*. SAGE, 2007.
- P. Mayring. *Qualitative Content Analysis. Theoretical Foundation, Basic Procedures and Software Solution*. Beltz, 2014.
- K. Petersen and C. Wohlin. Context in industrial software engineering research. In *Proceeding of the 3rd International Symposium on Empirical Software Engineering and Measurement*. IEEE, 2009.
- P. Runeson, M. Höst, A. Rainer, B. Regnell. *Case Study Research in Software Engineering. Guidelines and Examples*. Wiley, 2012.
- R. E. Stake. *The Art of Case Study Research*. SAGE, 1995.
- R. K. Yin. *Case Study Research. Design and Methods*. 3rd edition. SAGE, 2003.