CASE STUDIES

Research Method

Claudia Melo e Marco Aurélio Gerosa

IME-USP

04/2013

What is a case study?

- A case study is an empirical research method.
  - It is **not** a subset or variant of other methods, such as experiments, surveys or historical study.
- Best suited to applied problems that need to be studied **in context**
  - Phenomena under study cannot be separated from context. Effects can be wide-ranging.
  - **How** and **why** questions
- Settings where researcher has little control over variables, e.g. field sites.
- Effects take time to appear.
  - Days, weeks, months, or years rather than minutes or hours.

What is not a case study?

- Not an exemplar or case history
  - In medicine and law, patients or clients are "cases." A review of interesting instance(s) is called a case study.
  - Not a report of something interesting that was tried on a toy problem
- Not an experience report
  - Retrospective report on an experience (typically, industrial) with lessons learned
- Not a quasi-experiment with small n
  - Weaker form of experiment with a small sample size
  - Uses a different logic for designing the study and for generalizing from results

Why conduct a case study?

- **To gain a deep understanding of a phenomenon**
  - Example: To understand the capability of a new tool
  - Example: To identify factors affecting communication in code inspections
  - Example: To characterize the process of coming up to speed on a project
- **Objective of Investigation**
  - **Exploration-**To find what's out there
  - **Characterization-**To more fully describe
  - **Validation-**To find out whether a theory/hypothesis is true
- **Subject of Investigation**
  - An intervention, e.g. tool, technique, method, approach to design, implementation, or organizational structure
  - An existing thing or process, e.g. a team, releases, defects

When to use case studies

Strategy	Form of Research Question	Requires Control of Behavioral Events?	Focuses on contemporary events?
Experiment	How, why?	Yes	Yes
Survey	Who, what, where, how many, how much?	No	Yes
Archival Analysis	Who, what where, how many, how much?	No	Yes/No
History	How, why?	No	No
Case Study	How, why?	No	Yes

How can I tell it's a case study?

- Has research questions set out from the beginning of the study
- Data is collected in a planned and consistent manner
- Inferences are made from the data to answer the research questions
- Produces an explanation, description, or causal analysis of a phenomenon
  - Can also be exploratory

What is a case study?

- A case study is an empirical inquiry that
  - Investigates a contemporary phenomenon within its real-life context, especially when
  - The boundaries between phenomenon and context are not clearly evident.
- The case study inquiry
  - Copes with the technically distinctive situation in which there will be many more variables of interest that data points, and as one result
  - Relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result
  - Benefits from the prior development of theoretical propositions to guide data collection and analysis.

Parts of a Case Study Research Design

- A research design is a "blueprint" for a study
  - Deals more with the logic of the study than the logistics
  - Plan for moving from questions to answers
  - Ensures that the data is collected and analyzed to produce an answer to the initial research question
  - Strong similarities between a research design and a system design
- Five parts of a case study research design
  - Research questions
  - Propositions (if any)
  - Unit(s) of analysis
  - Logic linking the data to the propositions
  - Criteria for interpreting the findings

Part 1: Study Questions

- Case studies are most appropriate for research questions that are of the "how" and "why" variety
- The initial task is to clarify precisely the nature of the study questions (i.e. make sure they are actually "how" or "why" questions)
- Examples:
  - "Why do 2 organizations have a collaborative relationship?"
  - "Why do developers prefer this tool/model/notation?"
  - "How are inspections carried out in practice?"
  - "How does agile development work in practice?"
  - "Why do programmers fail to document their code?"
  - "How does software evolve over time?"
  - "Why have formal methods not been adopted widely for safety critical applications?"
  - "How does a company identify which software development projects to start?"

Types of Case Studies

- **Explanatory**
  - Judge between competing explanations
  - Example: How important is implementation bias in requirements engineering?
    - Rival theories: existing architectures are useful for anchoring, vs. existing architectures are over-constraining during RE

Types of Case Studies

- **Explanatory**
  - Adjudicates between competing explanations
  - Example: How important is implementation bias in requirements engineering?
    - 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 do software immigrants naturalize?

Types of Case Studies

- **Explanatory**
  - Adjudicates between competing explanations
  - Example: How important is implementation bias in requirements engineering?
    - 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 do software immigrants naturalize?
- **Causal**
  - Looks for causal relationship between concepts
  - Example: Requirements errors are more likely to cause safety-related defects than programming errors are
    - See study by Robyn Lutz on the Voyager and Galileo spacecraft

Types of Case Studies

- **Explanatory**
  - Adjudicates between competing explanations
  - Example: How important is implementation bias in requirements engineering?
    - 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 do software immigrants naturalize?
- **Causal**
  - Looks for causal relationship between concepts
  - Example: Requirements errors are more likely to cause safety-related defects than programming errors are
    - See study by Robyn Lutz on the Voyager and Galileo spacecraft
- **Exploratory**
  - Seeking new insights, generating ideas and hypotheses for new research.
  - Example: Christopher Columbus' voyage to the new world
  - Example: What do CMM level 3 organizations have in common?

Part 2: Study Propositions

- Propositions are statements that help direct attention to something that should be examined in the case study, i.e. point to what should be studied

  - Example: "Organizations collaborate because they derive mutual benefits"

- Propositions will tell you where to look for relevant evidence

  - Example: Define and ascertain the specific benefits to each organization

- Some studies may not have propositions –this implies a topic of "exploration"

  - Note: Even exploratory studies should have both <u>clearly-stated purposes</u> and <u>clearly-stated criteria for success</u>

Part 3: Unit of Analysis

- The unit of analysis defines <u>what a "case" is</u> in a case study
  - Example: a unit of analysis (case) may be an individual, and the case study may be the life history of that person
- Other units of analysis include decisions, social programs, processes, changes
  - Note: It is important to clarify the definition of these cases as they may be subjective, e.g. the beginning and end points of a process

Part 3: Unit of Analysis

- What unit of analysis to use generally depends on the primary research questions

- Once defined, the unit of analysis can still be changed if desired, as a result of discoveries based on data

- To compare results with previous studies (or allow others to compare results with yours), try to select a unit of analysis that is or can be used by others

Examples of Units of Analysis

- For a study of how software immigrants naturalize
  - Individuals
  - Development team
  - Organization
- For a study of pair programming
  - Programming episode
  - Pairs of programmers
  - Development team
  - Organization
- For a study of software evolution
  - Modification report
  - File
  - System
  - Release
  - Stable release
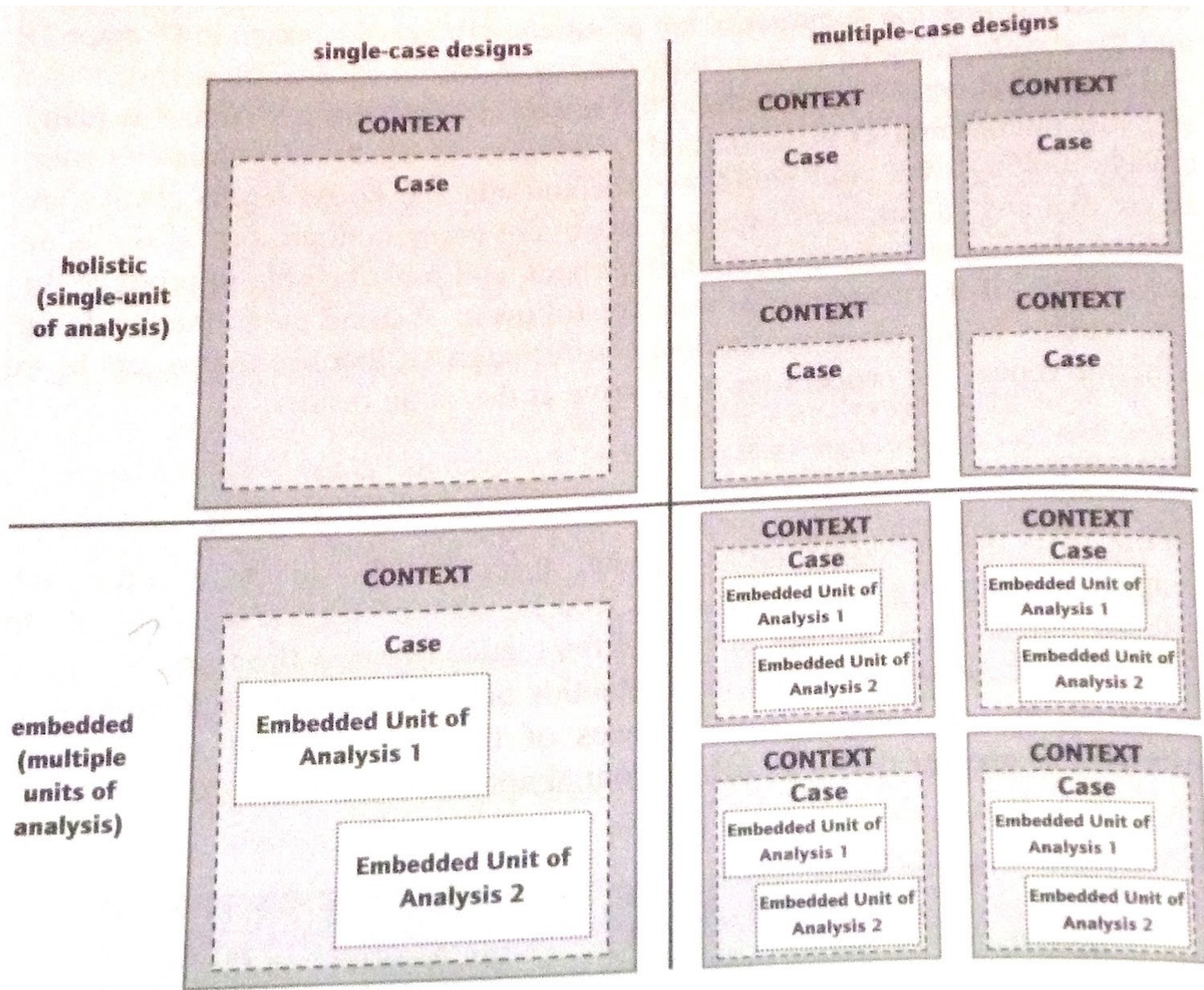
Multiple sources of evidence
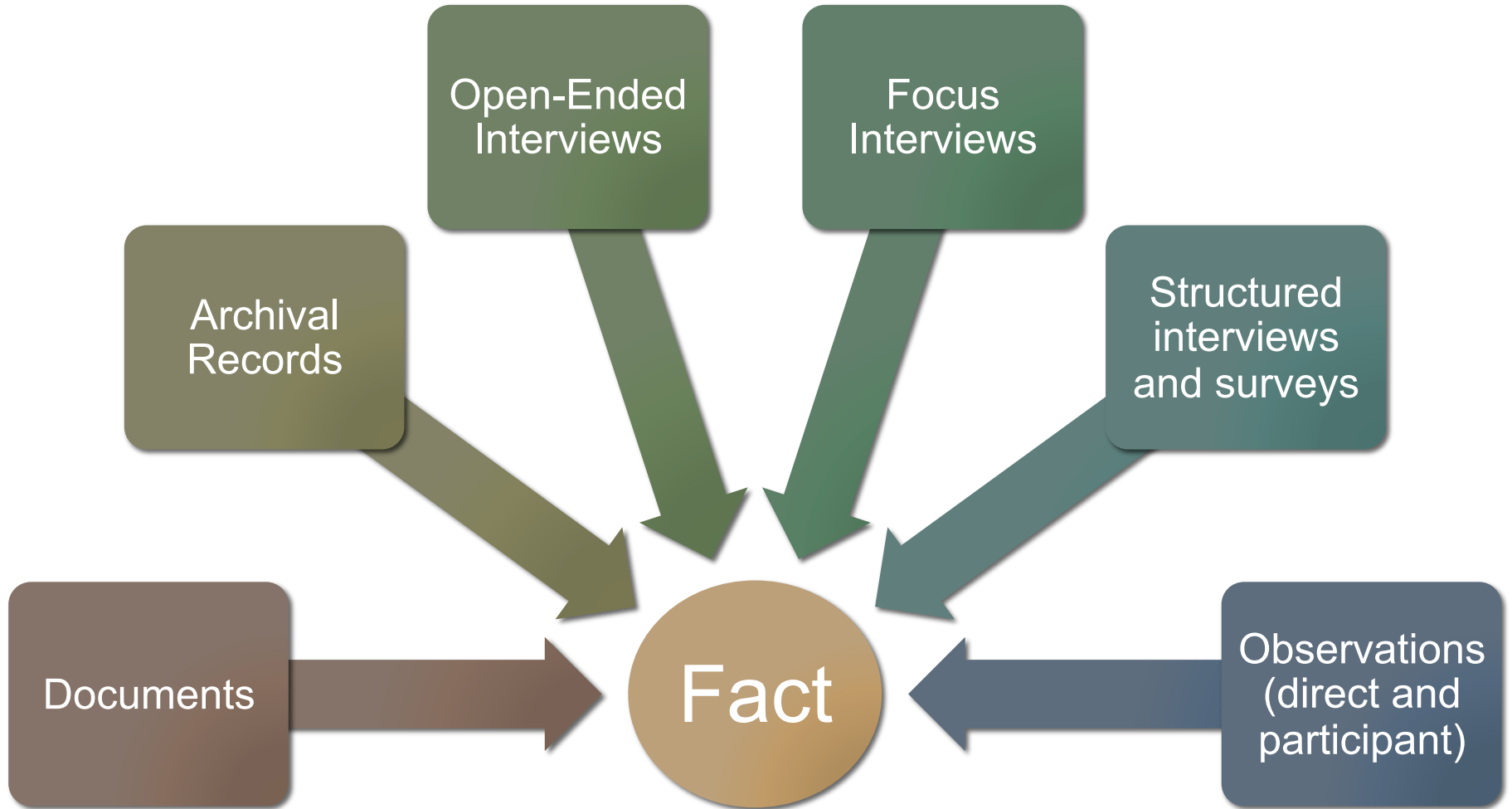
Triangulating

- 4 types of triangulation:
  - *of data sources* (**data triangulation**)
  - *among different evaluators* (**investigator triangulation**)
  - *of perspectives to the same data set* (**theory triangulation**)
  - *of methods* (**methodological triangulation**)

Part 4: Case Study Design

Holistic or Embedded

	single-case designs	multiple-case designs
holistic (single-unit of analysis)	CONTEXT — Case	CONTEXT — Case / CONTEXT — Case / CONTEXT — Case / CONTEXT — Case
embedded (multiple units of analysis)	CONTEXT — Case — Embedded Unit of Analysis 1, Embedded Unit of Analysis 2	CONTEXT — Case — Embedded Unit of Analysis 1, Embedded Unit of Analysis 2 (×4)

Holistic Designs

- **Strengths**
- Advantageous when no logical subunit can be identified or when the relevant theory underlying the case study is itself a holistic nature
- **Weaknesses**
- The entire case study may be conducted at an unduly abstract level, lacking sufficiently clear measures of data

Embedded Designs

- **Strengths**
- Introduces higher sensitivity to "slippage" from the original research questions
- **Weaknesses**
- Can lead to focusing only on the subunit (i.e. a multiple-case study of the subunits) and failure to return to the larger unit of analysis

Multiple-Case Designs

- If the same study contains more than a single case, it is a multiple-case design
- **Advantages**
- Evidence is considered more compelling
- Overall study is therefore regarded as more robust
- **Disadvantages**
- Rationale for single-case designs usually cannot be satisfied by multiple cases
- Can require extensive resources and time

Multiple-Case Designs: Holistic or Embedded

- A multiple-case study can consist of multiple holistic cases or multiple embedded cases, depending on the type of phenomenon being studied and the research questions

- Note there is no mixing of embedded and holistic cases in the same multiple-case study

- It is also important to note that for embedded studies, subunit data is NOT pooled across the subunits, but is used to draw conclusions for the subunit's case only

Selecting Case Study Designs: Single or Multiple?

- If you have a choice and the resources, multiple-case designs are preferred

- Single-case studies are often criticized due to fears about uniqueness surrounding the case

Selecting Case Study Designs - Closed/Flexible?

- A case study's design can be modified by new information or discovery during data collection

- If you modify your design, be careful to understand the nature of the alteration:

  - Are you merely selecting different cases, or are you also changing the original theoretical concerns and objectives?

  - Flexibility in design does not allow for lack of rigor in design

Part 5: Replication in Multiple-Case Studies

- When using multiple-case studies, each case must be carefully selected so that it either:
  - Predicts similar results (literal replication)
  - Predicts contrasting results but for predictable reasons (theoretical replication)
- With replication procedures, a theoretical framework must be developed that states the conditions under which a particular phenomenon is likely to be found (a literal replication) and the conditions when it is not likely to be found (a theoretical replication)
  - This framework is used to generalize to new cases

Replication Approach for Multiple-Case Studies

Yin, 2009. Pg. 57
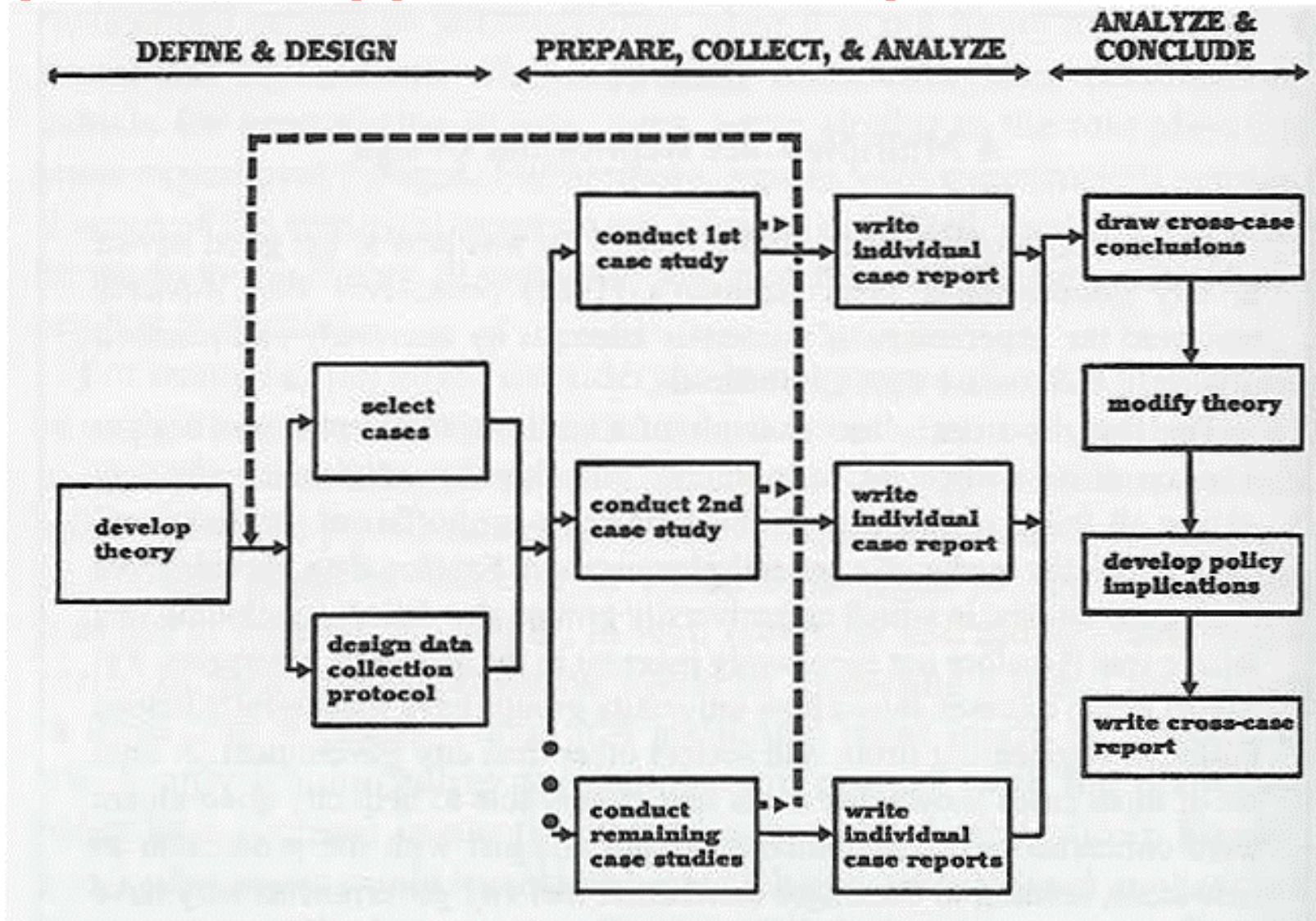
Rationale for Multiple-Case Designs

- Multiple-case designs are useful when literal or theoretical replications would provide valuable information for the study

- More results that back your theory typically adds more credibility to your case study

Part 6: Data Analysis & Interpretation

- Analytic Strategies
- 3 general strategies
- 5 specific analytic techniques
- Criteria for high quality analysis

Characteristics of Case Study Analysis

- Data analysis consists of examining, categorizing, tabulating, testing and recombining both quantitative and qualitative evidence to address the initial propositions of a study

- Analyzing case study evidence is difficult because strategies and techniques have not been well defined

- Every case study should have a general analytic strategy to define priorities for what to analyze and why

Criteria for High Quality Analysis

- Present all the evidence
- Develop rival hypotheses
- Address all major rival interpretations
- Address most significant aspect of the case study
- Use prior or expert knowledge

Objectives of Analytical Study

- Produce high quality analyses
- Present all evidence and separate them from any interpretation
- Explore alternative interpretations

Needs for Analytic Strategies

- Investigations on how the evidence is to be analyzed easily become stalled
- Analytic tools can only be helpful if the investigators know what to look for
- Analytic strategies are needed to address the entire case study since verbatim and documentary texts are usually the initial phase

Benefits of Analytic Strategies

- Put the evidence in preliminary order and treat the evidence fairly

- Prevent false starts

- Save time

- Produce compelling analytic conclusions

- Rule out alternative interpretations

- Help investigators use tools and make manipulations effectively

Three General Strategies

- GS1. Relying on Theoretical Propositions
- GS2. Thinking about Rival Explanations
- GS3. Developing a Case Description

GS1- Relying on Theoretical Propositions

- Shapes the data collection plan and gives priorities to the relevant analytic strategies

- Helps to focus attention on certain data and to ignore other useless data

- Helps to organize the entire case study and define alternative explanations to be examined

GS2 -Thinking About Rival Explanations

- Defines and tests rival explanations
- Relates to theoretical propositions, which contain rival hypotheses
- Attempts to collect evidence about other possible influences
- The more rivals the analysis addresses and rejects, the more confidence can be placed in the findings

GS3 - Developing a Case Description

- Analytic strategy that develops a descriptive framework for organizing the case study

- Serves as an alternative when theoretical proposition and rival explanation are not applicable

- The structure can reflect main aspects of the object of research or the general characteristics and relations of the phenomenon in question.

Five Specific Analytic Techniques

- Pattern Matching

- Explanation Building

- Time-Series Analysis

- Logic Models

- Cross-Case Synthesis

- *Note: They are intended to deal with problems of developing internal and external validity in doing case studies*

AT 1 - Pattern Matching

- Pattern matching compares an empirically based pattern with a predicted one

- If the patterns coincide, the results can strengthen the internal validity of the case study

- This is done by systematically searching for patterns (a) between categories or (b) dimensions or (c) data sources.

AT 2 - Explanation Building

- Analyzes the case study data by building an explanation about the case

- Stipulates a presumed set of causal links, which are similar to the independent variables in the use of rival explanations

- Has mostly occurred in narrative form

- May lead to starting a cross-case analysis, not just an analysis of each individual case

- Disadvantage: may drift away from original focus

AT 2 - Explanation Building

- Series of iterations in building explanation
  - Making initial theoretical statement
  - Comparing the findings of the initial case against such a statement
  - Revising the statement
  - Comparing other details of the case against the revision
  - Comparing the revisions to the facts of 2nd, 3rd or more cases
  - Repeating the process if needed

AT 3 - Time Series Analysis

- In *time-series analysis*, several points in time are investigated in order to detect changes in patterns over time. The data can then be used to identify a significant trend. It is often easier to look at one variable at a time.

- Three types of Time Series Analyses:
- Simple Time Series
- Complex Time Series
- Chronologies

Time Series Analysis: Simple Time Series

- Trace changes over time
- Single variable only, so statistical analysis of data is possible
- Match between a trend of data points compared to
  - significant trend specified before investigation
  - rival trend specified earlier
  - any other trend based on some artifact or threat to internal validity

Time Series Analysis: Complex Time Series

- Contain multiple set of variables (mixed patterns) which are relevant to the case study
- Each variable is predicted to have different pattern over time
- Create greater problems for data collection, but lead to elaborate trend that strengthens the analysis
- Any match of a predicted with an actual time series will produce strong evidence for an initial theoretical proposition

Time Series Analysis: Chronologies

- Trace events over time
- Sequence of a cause and effect cannot be inverted
- Some events must be followed by other events on a contingency basis after an interval of time
- Cover many different types of variables
- Goal is to compare chronology with that predicted by the explanatory theory

AT 4 - Logic Models

- Stipulate a complex chain of events over time
- Try to detect cause-effect-cause-effect patterns
- Match empirically observed events to theoretically predicted events
- Four types of logic models:
  - Individual-Level Logic Model
  - Firm or Organizational-Level Logic Model
  - An alternative configuration for an Organizational-Level Logic Model
  - Program-Level Logic Model

AT 5 -Cross-Case Synthesis

- Case study consists of at least 2 cases
- Using multiple case studies will
    - Treat each individual case study as a separate study
    - Have to create word tables that display data from individual cases according to some uniform framework
    - Examine word tables for cross-case patterns
    - Rely strongly on argumentative interpretation, not numeric properties

Part 7: Generalizing from Case Study to Theory

- Theory for case studies is characterized as *analytic generalization* and is contrasted with another way of generalizing results known as *statistical generalization.*

- In *analytic generalization*, previously developed theory is used as a template against which to compare the empirical results of the case study.

Reductionist fallacy

- "No man ever steps in the same river twice, for it's not the same river and he's not the same man" Heraclitus

Naturalistic generalization

"Is a process where readers gain insight by reflecting on the details and descriptions presented in case studies. As readers recognize similarities in case study details and find descriptions that resonate with their own experiences, they consider whether their situations are similar enough to warrant generalizations"

Encyclopedia of Case Study Research

How can I evaluate a case study?

- Using the same criteria for other empirical research
- **Construct Validity**
- Concepts being studied are operationalized and measured correctly
- **Internal Validity**
- Establish a causal relationship and distinguish spurious relationships
- **External Validity**
- Establish the domain to which a study's findings can be generalized
- **Experimental Reliability**
- Demonstrate that the study can be repeated with the same results

What Makes an Exemplary Case Study?

- The exemplary case study goes beyond the methodological procedures
- Mastering the techniques does not guarantee an exemplary case study

Characteristics of an Exemplary Case Study

- 1. The Case Study Must Be Significant
  - The case should be unusual and of general public interest
  - The issue are nationally important, either in theory or practical terms
  - Prior to selecting a case study, the contribution should be described in detail assuming that the intended case study were to be completed successfully
- 2. The Case Study Must be "Complete"
  - Completeness can be characterized in at least three ways:
    - The boundaries of the case are given explicit attention
    - Exhaustive effort is spent on collecting all the relevant evidence
    - The case study was not ended because of nonresearch constraints

Characteristics of an Exemplary Case Study

- 3. The Case Study Must Consider Alternative Perspectives
  - The case study should include consideration of rival propositions and the analysis of the evidence in terms of such rivals
  - This can avoid the appearance of a one-sided case
- 4. The Case Study Must Display Sufficient Evidence
  - The report should include the most relevant evidence so the reader can reach an independent judgment regarding the merits of the analysis
  - The evidence should be able to convince the reader that the investigator "knows" his or her subject
  - The investigator should also show the validity of the evidence being presented
- 5. The Case Study Must Be Composed in an Engaging Manner
  - A written case study report should be able to entices the reader to continue reading

Case Study as a Research Method

- The case study is a distinct research method with its own research designs
  - It is not a subset or variant of research designs used for other strategies (such as experiments)
- Scientific
  - Synergistic relationship between theory and data
  - Starting a case study requires a theoretical orientation, which drives data collection
- Useful for answering "how" and "why" questions
  - In contrast to who, what, when, how many, how much
  - How, why = explanatory, descriptive
- Does not require control over events
  - More observational
- Focus on contemporary events
  - Less historical

More resources

- Case study protocol
- www.dur.ac.uk/ebse/resources/templates/CaseStudyTemplate.pdf

References

- Perry, Sim & Easterbrook, Case Studies for Software Engineering, ICSE 2004 Tutorial
- Robert Yin (2009). Case study research. Design and methods. 4th edition.