

## Evidence-Based Software Engineering: Case Studies

### What is a Case Study?

According to (Yin, 2003) a *case study*<sup>1</sup> is an empirical enquiry that:

- investigates a contemporary phenomenon within its real-life context (an experiment removes it from its context, and while a quasi-experiment may preserve context, it does so at the expense of having little control);
- is especially appropriate when the boundaries between the phenomenon being studied and its context are not clearly evident, and so may not be easily distinguished.

And use of a case study:

- copes with the technically distinctive situation where there will be many more variables of interest than data points;
- relies upon multiple sources of evidence, using *triangulation* of the data for its analysis;
- benefits from the use of theoretical *propositions* to guide data collection and analysis.

Hence case studies are particularly appropriate as a form of empirical study in those situations:

- when the subject being studied cannot easily be divorced from its real-world context;
- when the interest lies in the *interconnection* of factors, rather than in trying to isolate and control them, and to study them separately (as we might do in a quasi-experiment);
- where there may be many sources and forms of data that can be drawn upon;
- where we need a *deeper* study based upon using more than one instance;
- for *longitudinal* studies where a case spans a lengthy period of time (months, years,...).

So where experiments and surveys provide statistical rigour, case studies emphasis depth.

### What a Case Study is *not*...

...A case study is not a *narrative* – the term ‘case study’ is used for teaching subjects such as business, where students may study a particular instance, and the narrative description is termed a case study. However, this is a quite different use of the term. Also, it is not a *scenario* – a meaning sometimes used in computing where a researcher is describing a particular scenario to which they have applied a modelling technique, method, or program. Such a scenario may simplify a real-life situation, whereas **case study research** is concerned with the complexities of the real world.

### Types of Case Study

Yin (2014) categorises case studies into three types: explanatory; descriptive and exploratory.

- an *explanatory* study is used to answer question about *how* some phenomenon works and *why* it works, in situations where the researcher may have little or no control of the phenomenon but may be able to draw upon different theories;
- *descriptive* case studies are used to provide a rich and detailed analysis of a phenomenon and its context—telling a story including different perceptions, but less study in detail about mechanisms;
- finally, an *exploratory* study is used to help define the questions or hypotheses needed for a subsequent study, or to help understand a problem.

### Organisation of Case Studies

An important distinction for case studies lies between:

- the **single-case** form, which is: less trustworthy than multiple case forms; appropriate where there is a critical case to consider; appropriate where we want to study an extreme case; or even for studying a representative case.

---

<sup>1</sup> We should note that Yin’s view of case study research is *positivist*. A positivist approach is one where we believe that there are general laws, construct models and theories, and then test them in some way. This philosophy underlies the scientific method and emphasises objectivity, with the researcher seeking to be neutral and impartial. In contrast, the *interpretivist* philosophy believes that phenomena are only relevant within a context and that there are multiple realities accessed through social constructs such as language.

- the **multiple-case** form, which: provides more compelling evidence; makes it possible to use *replication logic*, whereby different cases predict the same results (or different ones, if there are good reasons and different results are predicted by theory).

In addition, there is a further choice to make between:

- a *holistic* form (one unit of analysis)
- an *embedded* form (where multiple units are analysed at different levels)

### Designing a Case Study

As defined by Yin (2014), this involves five major steps:

1. *Determine the study questions*: involves identifying the **research question** that represents the high level concern of the study (e.g. “investigate changes to the development process that arise from adopting agile methods in an organisation”)
2. *Identify any propositions*: where the propositions are more detailed than the research question and identify specific issues to investigate, for example, that agile methods should:
  - reduce development time
  - increase customer satisfaction
  - improve developer motivation
3. *Select the unit(s) of analysis*: this involves two major elements:
  - determining what a **case** is, such as: a typical instance; an extreme instance; an instance that has elements that will help test out a theory; a convenient instance
  - deciding upon the form of **unit** to use: company; development project; technology.
4. *Determine the logic that links the data to the propositions*: here the design task is one of determining what data needs to be collected in order to evaluate the propositions (for example, “increase customer satisfaction” might involved a conducting structured interviews with questions about past and current satisfaction, and where the different dimensions of ‘satisfaction’ need to be identified along with ways of measuring them). There is a related issue in the role of the researcher in data *collection*, either an *observer* or as a *participant*. This in turn determines what data can be collected.
5. Define the criteria to be used for interpreting the findings: a task that is concerned with analysing and evaluating the data to answer such questions as how well the findings support the proposition; what ‘level’ of satisfaction is considered to provide support; and whether the findings are clear enough to allow other explanations to be rejected.

### Generalising from Case Studies

This is one of the limitations that is encountered when using case studies—strictly, we cannot generalise from a single case (however, if we have chosen a ‘typical’ case, we might be able to generalise to similar situations). Because the data and the subsequent analysis are largely qualitative, it is not possible to use statistical arguments to aid generalisation of the results.

For software engineering, case studies are often the only way we can perform ‘field trials’ using real-world situations. There are few situations where we can perform more rigorous studies, and hence many of our field studies are qualitative rather than quantitative.

### Case Studies and Theory

A case study can be used to:

- *build a new theory*: which itself might be a conceptual framework that can then be used to analyse further case studies, or new hypotheses than can be tested by formal experiments or surveys;
- *test an existing theory*: by providing empirical support that supports or contradicts the theory;
- *evaluate between alternative theories*: by matching competing theories best to the outcomes of the case study.

(For computing, the idea of a ‘theory’ is equivalent to that of a ‘model’ for many purposes, for example, a particular form of agile development. With case studies being used to determine when it is best used.)