# Theory of Science and Research Methodology

**Rolf Johansson**
rolfj@infra.kth.se
**Associate Professor**
Department of Infrastructure
Urban Studies / Built Environment Analysis
Royal Institute of Technology, Stockholm

Mars 2004
Revised September 2004

# Theory of Science

## THE EMPIRICAL SCIENCES

Scientific research may be divided into two groups:

- Empirical, factual or descriptive sciences, and
- non-empirical, formal or theoretical sciences.

The empirical sciences aim at exploring, describing, understanding or explaining, and predicting events in *the world we are living in*. Statements in the empirical sciences must be controlled by facts from experience, which are empirical facts. Within non-empirical sciences (that is, the mathematical—logical sciences) statements need not refer to empirical facts. Here we will mainly discuss the empirical sciences.

We might say that scientific inquiry in the empirical sciences has two main components:

- ascertaining and discovering facts (manufacturing or finding evidence), and
- constructing hypotheses and theories (theory construction).

### Evidence

Research is about creating new knowledge, and the raw material of research is evidence. When the world we live in is the subject matter of inquiry, empirical data are collected. Empirical facts are collected in various manners, for instance,

- through experiments,
- by observation,
- through interviews,
- by surveys,
- by psychological or clinical testing, and
- through artefacts and documents.

In an experiment, evidence is actually manufactured in a laboratory setting: evidence is a product of the experiment. In a naturalistic inquiry evidence is there, already existing, and has to be uncovered and tested.

### Theory construction

Theory construction can be said to serve two main purposes:

(1) to predict the effects of particular experiments, and
(2) to make intelligible previously known facts.

Both purposes are present in the natural sciences. In the social sciences the second purpose is dominating and in the humanities only the second purpose is operative. In the natural sciences cause and effect relationships are established, and in the humanities the purpose of theory–construction is normally to conceptualise phenomena.

Problems concerning the interrelation of the various concepts mentioned here—description, explanation, prediction, and theory—might be considered in the light of intellectual history.

**Two traditions in scientific thinking**
There are two main traditions in scientific thinking: positivism and hermeneutics. These can be distinguished in the history of ideas, differing as to the conditions an explanation has to satisfy in order to be scientifically respectable. The contrast between the two traditions is usually characterised as causal versus teleological explanation (or explanation of intention).

Positivism is the philosophical foundation of the natural sciences. The natural sciences, in the modern sense, were born in the 16th and 17th centuries. In the Middle Ages the works of ancient philosophers and the Bible where regarded as the main sources of knowledge. The authority of the ancient philosophers was questioned by the natural scientific revolution, and the statements made by these philosophers tested empirically by experiments. The philosopher Francis Bacon (1561—1626) argued that to understand nature, we have to study nature—not the writings of Aristotle.

Auguste Comte (1798—1857) and John Stuart Mill (1801—1893) typically represent positivism within the philosophy of science. According to von Wright (1971) the characteristic features of positivism are:

- *Methodological monism.* The idea that there is one scientific method that can be applied within all different fields of scientific investigation.
- *The natural sciences as an ideal.* The view that the exact natural sciences, in particular mathematical physics, set a methodological ideal or standard which apply to all the other sciences, including the humanities.
- *General explanatory laws.* Explanations are causal and consist in the subsumption of individual cases under hypothetically assumed general laws of nature, including even “human nature”.

The revolution in the natural sciences during the Renaissance and the Baroque era was to a certain extent paralleled by the formation of the humanities. The humanities appeared on the intellectual stage with a scientific claim early in the nineteenth century. Man, his history, languages, mores, and social institutions, were studied systematically. This was the awakening of the humanities and their philosophical foundation: hermeneutics.

Since the humanities were newcomers and the natural sciences had already been established for several hundred years, methodology and the philosophy of science became concerned mostly with the relationship between the two main traditions of empirical inquiry: positivism and hermeneutics.

The anti–positivist philosophy of science—hermeneutics—is a much more diverse and heterogeneous trend than positivism. Some of the best known advocates of hermeneutics are Johann Gustav Droysen (1808–1884), Wilhelm Dilthey (1833–1911), Georg Simmel (1858–1918), Wilhelm Windelband (1848–1915), and Robin George Collingwood (1889–1943). All these thinkers reject the methodological monism of positivism and refuse to view the exact natural sciences as the only ideal for a rational understanding of reality.

Many anti–positivist thinkers emphasise the contrast between the natural sciences and the humanities. On the one hand are the sciences such as physics, chemistry, or physiology, which aim at generalisations about reproducible and predictable phenomena. On the other hand are sciences that, like history, aim to grasp the individual and unique features of their objects. The German philosopher Windelband made a distinction between "nomothetic" sciences and "idiographic" sciences. The former search for general explanatory laws, and the latter aim at a descriptive study of individuality. Windelband (1915) claims:

> In creating knowledge of reality, the empirical sciences search for either the general in the shape of a natural law, or the individual in the historically decided shape; some look for the form, which is always the same; some for the within itself totally specific content of the true process which once existed. Some are sciences of law, other of event. The former teach that what always is, the latter that which once was. Scientific thought is … in one case nomothetic, in the other idiographic. (Translated by the author).

Nomothetic thinking seeks explanations through general laws, and idiographic thinking seeks understanding of specific events. But, as Windelband emphasises, the same object may be investigated with both nomothetic and idiographic methods. The behaviour of a human being, for instance, may be investigated from the perspective from which we are all alike, or from an interest in understanding human actions in specific, unique, situations.

The anti–positivists also attacked the accepted view of explanation. The German philosopher Johann Gustav Droysen introduced a dichotomy between *explanation* and *understanding* (*Erklärung* and *Verstand*). The aim of the natural sciences, he argued, is to explain; the aim of history is to understand the phenomena that fall into its domain. Wilhelm Dilthey systematically developed these ideas. For the entire domain of "the understanding method" he used the name *Geisteswissenschaften.* "Moral sciences" is probably the best translation into English.

It is difficult to make a sharp distinction between the words "explain" and "understand". Explanation can be said to further our understanding of things. But "understanding" also has a psychological ring which "explanation" has not.

This psychological feature was emphasised by several of the nineteenth-century anti–positivist methodologists. Georg Simmel was one of these. He thought that understanding, as a method characteristic of the humanities, is a form of empathy (*Einfühlung*). The scholar has to, in his own mind, recreate the mental atmosphere—the thoughts and feelings and motivations—of the objects of his study.

Understanding is also connected with intentionality in a way explanation is not. One understands the aims and purposes of an agent, the meaning of a sign or a symbol, and the significance of a social institution or religious rite. The British archaeologist and philosopher Robin George Collingwood emphasised this aspect of understanding by contrasting natural and historical processes. Collingwood (1946/1994) argued that historical processes are unique, but processes in nature are repeated. Therefore, history proper is the history of thought. What are called "events" in history are really actions that express some thought (intention, purpose) of its agent.

If one accepts a fundamental methodological cleavage between the natural sciences and the historical sciences, a question arises about where the social and behavioural sciences stand. These sciences were born largely under the influence of both positivist and anti–positivist tendencies in the last century. The social sciences have become a battleground for the two opposed trends in the philosophy of scientific method.

The heyday of positivism in the middle of the nineteenth century was succeeded by an anti–positivist reaction towards the end and around the turn of the century. But in the decades between the wars positivism returned. The new movement was called neo-positivism or logical positivism, later also logical empiricism.

Social scientific methodology is very diverse. When the social sciences were singled out as a specific branch within the humanities at the beginning of the twentieth century, both qualitative case study methods from the hermeneutic tradition and quantitative statistical and experimental methods from the natural sciences were practiced. Later, during the period of logical positivism in philosophy between the wars, quantitative methods dominated within the social sciences. In 1958, in his essay *The Idea of a Social Science and its Relation to Philosophy*, the philosopher Peter Winch attacked positivism and defended an understanding of social phenomena by methods different in principle from those of the natural sciences. Since then there have been attempts to merge quantitative and qualitative research within the social sciences, and for the last twenty years a second generation of qualitative case study methodology has developed.

**Main sources and suggested further reading:**
Georg v. Wright (1971) and Peter Winch (1958/1990).

## CONCEPTUAL FOUNDATIONS OF RESEARCH

Scientific knowledge is dependent on both reason and experience. Scientists may work on a conceptual–theoretical level, or an observational–empirical level, or both. I will now focus on the conceptual–theoretical level.

The most important achievements within the sciences are not inventions and discoveries, but the development of new concepts and theories. New concepts change human relations in the cultures in which they are introduced. We see the world through our concepts. We interpret the world by *seeing* something *as* something.

### Concept

Science begins by developing concepts to describe the empirical world. A concept is an abstraction representing an object, a property of an object, or a certain phenomenon. Concepts do not actually exist as empirical phenomena: a concept is rather a *symbol* of the phenomenon. As a symbol of a phenomenon the concept functions as a means of communication.

Concepts also introduce a *perspective*: a way of looking at empirical phenomena. Through scientific concepts the perceptual world is given an order and coherence that could not be perceived before conceptualisation.

Another function of concepts is as a means of *classification* and *generalisation*. Scientists categorise, structure, order, and generalise their experiences and observations in terms of concepts. Concepts serve as components of theories and thus of explanations, predictions, and understanding.

There are different types of concepts: concepts of class, relation, and variable.

If phenomena that resemble each other in some respect are assembled in one group, we have created a concept of *class*. For instance all people with blue eyes fall under the concept of blue-eyed people.

Objects, individuals, or concepts of class may have a specific relation to each other. For instance, Stockholm and Dar–es–Salaam are geographically related to each other. Stockholm is north of Dar. "North of" is a concept of *relation*.

A concept of *variable* refers to something that is measurable. For instance the concept of "length" is a variable between different individuals.

**Theory**
Our concepts tend to create *systems of concepts*: that is, theories.

Nachmias & Nachmias (1992) classifies theories.
Theories can be classified according to their:
- scope: macro/micro theories,
- function: whether they deal with statics/dynamics, or structure/process,
- structure: logical interrelations/a loosely defined set of propositions, and
- level: low/high level of theorising.

Theories differ according to their level, from very simple classifications to advanced theoretical systems.

High level of theorising

| | |
|---|---|
| Theoretical Systems | A combination of taxonomies and conceptual frameworks |
| Conceptual framework | Descriptive categories in a structure that provide explanations and predictions |
| Taxonomies | A system of categories where the relations among categories can be described |
| Ad Hoc Classificatory System | Empirical observations organised and summarised in categories |

Low level of theorising

**Model**
Closely related to the idea of theory as a systematic conceptual organisation is the notion of a model. A model can be viewed as a likeness of something.

The drawings of a building are a kind of model. They are a visual representation of the structure and features of the real building. A model may also be a scaled-down version of a physical artefact.

In the social sciences, models usually consist of symbols rather than physical matter: that is, the characteristics of some empirical phenomena, including their components and the relationships between the components, are represented in logical arrangements among concepts.

A model is an abstraction from reality that serves the purpose of ordering and simplifying our view of reality while still representing its essential characteristics.

## Definition

Concepts have to be clear and precise, and this is achieved by definitions.

A *lexicalic* (or *descriptive*) definition describes the meaning of a term in linguistic use; for instance, “The term T usually means ...” An *ostensive definition* points out what the expression is used for; for instance, “The colour red is *this* colour.” These two kinds of definitions are seldom used in science.

A *stipulative* (or *conceptual*) definition prescribes that a term will be used in a certain context with a certain meaning; for instance, “In this study ‘*modern*’ refers to the period stretching from 1930 until today.” This kind of definition is probably the most frequent in scientific disciplines.

An *operational definition* lays down the meaning of a term by referring to a certain operation or method that shows how to measure or show the presence, degree, or size of what the expression stands for. Operational definitions are common in the sciences, specifically in the natural sciences.

An *enumerative definition* explains the meaning of a term by enumerating the phenomena that can be ranged within the term; for instance, by fruit I mean: apple, banana, orange ... Enumerative definitions are seldom used in the sciences.

According to Nachmias & Nachmias (1992) there are four criteria that a good conceptual definition must fulfil:

- A definition should *not be circular*. An example: “*User-oriented design* refers to design that is oriented towards the user” is a circular definition. Sometimes the circularity is not that obvious, for instance, when a term A is defined by a term B, and term A is subsequently used to explain term B.
- A definition must point out the unique attributes or qualities of whatever is defined. It must be inclusive of all cases it covers and exclusive of all cases not covered by it.
- The definition should be *more easily understood* and clearer than the term being defined. The definition “*User-oriented design* is design that is good for the user”, for instance, is not very clear since there is much disagreement about what might be good for the user.
- A definition should be stated positively. “A is a property that lacks B” is not a good definition, since there may be a lot of other properties than A that lack B.

**Main sources and suggested further reading:**
**Nachmias & Nachmias (1992) and Lundequist (1999).**

## THE PRINCIPLES OF LOGIC

There are two established principles of logic: *deduction* and *induction*. According to the American pragmatist philosopher Charles Sanders Peirce, there is also a third principle of reasoning. Peirce calls it *abduction*. The result of deductive reasoning is a *fact* that is *definitely* true if the premises are true. The result of induction is a *principle* that is *probably* true. The result of abduction is a *case* that *may* be true.

### Deduction

In deductive logic there are two premises from which we draw a conclusion that must be true if both of the premises are true.

The chief business of the logician is to classify arguments. The classes are defined by certain typical forms of reasoning, called syllogisms. The most basic syllogism, for instance, is called *Barbara* and looks like this:

A is B
B is C
___
Hence, A is C

Or, to replace the letters with words:

All humans are mortal
This logician is a human
___
This logician is mortal

The result of the syllogism is true if and only if both premises are true. If one or both of the premises are false, the conclusion is false (or at least not certainly true).

This syllogism, for instance, is logically correct, but the conclusion is false. The reason is that one of the premises is false.

All humans are machines
This logician is a human
___
This logician is a machine

The *Barbara* syllogism is a brief schema for stating an argument from the general to the particular that consists of two statements and a conclusion that must be true if these two statements are true.

*Barbara* is, in fact, the application of a rule. The major premise lays down this rule, as, for example, does the premise 'all beans from this bag are white.' The other, or minor, premise states a case which falls under the rule, such as 'these beans are from this bag.' The conclusion applies the rule to the case and states a fact: 'these beans are white.'

> Rule: all the beans from this bag are white
> Case: these beans are from this bag
> ———
> Fact: these beans are white

All deduction is of this character: it is merely the application of general rules to particular cases. The result is a fact that *must* be.

**Induction**

Inductive reasoning is different: we conclude a rule from the observation of a fact in a certain case. Induction is the inference of the rule from a case and a fact.

> Case: these beans are from this bag
> Fact: these beans are white
> ———
> Rule: all the beans from this bag are white

With induction we generalise from a number of cases of which something is true, and infer that the same thing is true of a whole class, or, where we find a certain thing to be true of a certain proportion of cases, we infer that it is true of the same proportion of the whole class.

The inductive conclusion is a rule that reaches beyond the content of the premises. Therefore the conclusion may be false even if the premises are true. Inductive reasoning only supports a conclusion to some degree of probability. A characteristic feature of induction is that, based on what we actually know, we draw a conclusion about something that we do not know.

The inductive principle of reasoning can also be said to be based on repeated observation of facts within a case:

> This bean is from this bag, and it is white
> This second bean is from this bag, and it is white
> This third bean is from this bag, and it is white
> ...
> ———
> All beans from this bag are white

But, since we have not seen all the beans in the bag, our conclusion might be wrong: the next bean from the bag may be pink. There is obviously a problem with induction. This problem may be illustrated by a story inspired by Bertrand Russell (1912/1976): the story of the research-oriented turkey. During his first morning on the turkey farm, the turkey observed that food was provided at 9 o'clock. He had a scientific mind and did not want to draw hasty conclusions. He made more observations under different circumstances: different days of the week, different weather, and so on. Finally he was ready to make an inductive conclusion: “I always get food at 9 o’clock”. But it was a false conclusion. Next morning was Christmas, and he was beheaded and served for dinner himself.

Another problem of induction is that we cannot prove that the principle works without making a reference to the fact that it has worked before. Induction worked at one occasion, induction worked at another occasion ... induction always works. That is trying to prove induction by using induction.

**Abduction**
In Peirce’s words:

> Abduction is when we find some very curious circumstance, which would be explained by the supposition that it was a case of a certain general rule, and thereupon adopt that supposition. Or, where we find that in certain respects two objects have strong resemblance, and infer that they resemble one another in other respects. (Peirce 1931—35 & 1958/1992)

Abductive reasoning, if illustrated with the example of the white beans, looks like this:

Fact: these beans are white
Rule: all the beans from this bag are white
___
Case: these beans are from this bag

Abduction resembles induction insofar as the conclusion reaches beyond the content of the premises. But the difference between them is that the abductive conclusion identifies something of a new quality.

In the sciences the different modes of reasoning are combined. Abductive reasoning is active in creative acts; for instance, in formulating a hypothesis. Deduction is used to test a hypothesis: if an implication of a hypothesis in a specific case is theoretically deduced, and the case is then produced experimentally, and the theoretically deduced fact matches the observable fact, then the hypothesis is verified. Finally, induction is used to generate general statements or theories.

**Main sources and suggested further reading:**
Peirce (1931–35 & 1958 / 1992) and Rosing (1994).

## THE CASE OF IGNAZ SEMMELWEIS

The case of Semmelweis is often referred to in theory of science literature in connection with discussion of the hypothetic–deductive method (Hempel 1966, Rosing 1994).

I will use the case of Semmelweis' studies of childbed fever (puerperal fever) as an illustration of scientific research in practice.

Ignaz Semmelwies was a Hungarian doctor of medicine working at the first delivery ward in the public hospital in Vienna during the 1840s. He noticed that many women at the first delivery ward suffered from childbed fever, which is a serious and often deadly disease. About 10% of the patients died in childbed fever every year. In the second delivery ward only about 2% of the mothers died of the disease. Semmelweis was puzzled by this big difference between the wards. He had identified a problem, and that is the first step in the research process.

Semmelweis actually faced an ideal experimental (or quasi–experimental) situation. He had two wards that could be compared with each other. He could manipulate different variables in the first ward, study the effects, and compare them with the second ward to make sure that there was no other confounding variable over which he had no control. In a quasi–experimental design, ward one would be called an experimental group and ward two a control group.

This is a very common research design. It is used, for instance, when the effects of a new medicine is to be tested. Two groups of persons are established: an experimental group and a control group. The persons in the experimental group are given the new medicine and the persons in the control group are given a harmless substance (called a placebo). The effect of the medicine is measured by comparison between the two groups. If possible the two groups are also observed before testing to make sure that they do not differ in any respect.

We return to Semmelweis. To approach his research problem he could choose to use two different strategies: inductive reasoning or hypothetico–deductive reasoning. If Semmelweis had applied inductive reasoning he would have started to collect a lot of facts by observation. From these facts he would have inductively concluded a theory which would, hopefully, provide a solution to the problem.

But he used hypothetico–deductive reasoning (or, with reference to the previous chapter, abductive–deductive reasoning). He tried to make a guess about the cause of the high death rate in childbed fever in ward one, compared with ward two. This guess we call a hypothesis, and it must be tested deductively to prove it to be either true or false. To make testing possible, a statement (a test implication) must be deduced from the hypothesis about a fact that may be observed.

The hypothesis must fulfil two requirements:

- The hypothesis must produce *a possible solution to the problem*, and
- *it must be possible to test.*

The first step is to identify a test implication from the hypothesis. That is, the question ‘are there any effects that are possible to observe and that would occur if the hypothesis is true?’ must be answered. The observable effects are called test implications from the hypothesis.

A test implication is theoretically deduced from the hypothesis and expected to be present in a certain case. Then the case is produced, and if the test implication cannot be observed, the hypothesis must be false. But if the expected test implication is actually present, the hypothesis is verified.

When Semmelweis arrived at the hospital there were many attempts to explain the phenomenon of concentrated childbed fever in ward one. At that time it was believed that childbed fever might be epidemic. But if childbed fever were epidemic, then it would spread in whole areas. (Hypothesis [H]: childbed fever is epidemic. Test implication [I]: if H, then the disease would spread in whole areas). But, instead, Semmelweis could observe that the disease was much more frequent in ward one than in ward two. And furthermore, there were very few cases of childbed fever in Vienna outside the hospital. If childbed fever were epidemic like, for instance, cholera, it would not have been that selective. He also noticed that women that had a place in ward one and delivered their children on the way to the hospital very seldom caught childbed fever, even if they gave birth under very unfavourable circumstances. Semmelweis could not observe any expected implications from the hypothesis and concluded that it was false.

Another opinion was that ward one was overcrowded, but Semmelweis could show that ward two was actually more overcrowded than ward one.

That childbed fever is caused by diet was suggested as an explanation. If that were the case, then there must be a difference in the diet between the wards. Semmelweis observed that there where no differences, and accordingly, a difference in diet could not be the explanation he was looking for.

Now he turned to more challenging hypotheses. One rumour was that childbed fever was caused by fear, and that the patients were scared to death. In that case, there must be some circumstances in ward one that caused the patients to be more scared than those in ward two. And Semmelweis actually found that a priest that passed through the ward in order to give extreme unction to dying patients particularly frightened the patients in ward one. He came in full canonicals and was followed by a boy ringing a bell. He did not pass unnoticed.

Semmelweis now had a new hypothesis: childbed fever is caused by fear and the priest causes fear. His test implication was: if H is true, then if the priest is convinced to act discreetly, the death rate will fall. The priest acted differently, but Semmelweis could not observe any changes in the death rate. Hence, the hypothesis was false.

He tested several other hypotheses without any positive result. Then a colleague of Semmelweis'—doctor Kolletschka—cut his finger by mistake during an autopsy. Kolletschka soon got ill and suddenly died with symptoms very similar to those of childbed fever. Semmelweis realised that this could provide the answer to the riddle.

The medical students were examining patients in ward one, but not in ward two. The medical students also had to do autopsies as part of their education. They often came from an autopsy, washing their hands only with soap, and examined the patients in ward one. In ward two midwifes handled the patients, and they did not participate in autopsy. Semmelweis' new hypothesis was that the medical students transferred some substance that caused childbed fever from dead bodies to the patients. If H is true, Semmelweis concluded, then if the medical students wash their hands with disinfectant, childbed fever will decrease. Semmelweis ordered everyone that had done an autopsy to wash their hands with disinfectant afterwards, and he could observe that the death rate in ward one changed for the better.

This is an example of hypothetico–deductive reasoning in practice. Mario Bunge (1967) illustrates the hypothetico–deductive method in a model.

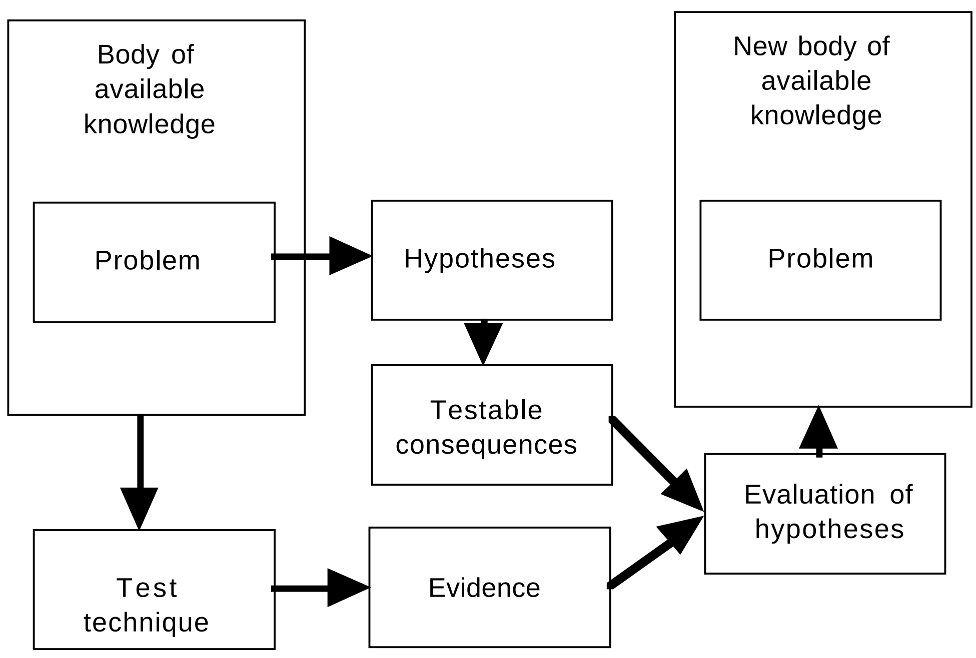



Figure 1. The research cycle as illustrated by Bunge (1967).

The starting point is a problem that is formulated within the body of available knowledge. A hypothesis that supplies a possible solution to the problem is formulated. Testable consequences are theoretically deduced from the hypothesis. These are the test implications. Some procedure to test the hypothesis in reality is designed and performed. The actual evidence produced is evaluated by comparing it with the test implications. If the expected evidence is present, then the hypothesis is verified and a solution to the problem is at hand. We now have a new body of knowledge and new research problems.

**Karl Popper and falsificationism**

The philosopher Karl Popper (1972) argues that a scientific statement should not be defined as a statement that can be verified. Instead it should be defined as a statement that runs the risk of being *falsified*. According to Popper, the scientist has to focus on circumstances in which a hypothesis might not hold, instead of looking for evidence that supports the hypothesis. A hypothesis is true only as long as it has not been falsified. Every theory and hypothesis is thus provisional, and will sooner or later be replaced by a better provisional theory.

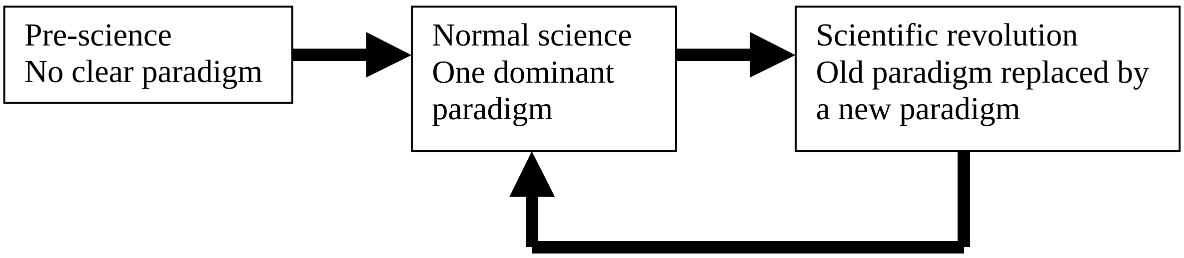
**Thomas Kuhn and the role of paradigm**

In his book *The Structure of Scientific Revolution*, Thomas Kuhn (1962) observes from the history of science that scientists do not work in the way Popper suggests. One could of course argue that they should change and actually follow Poppers advice and focus on falsification. But science has been rather successful, so maybe there is no reason to change.

Kuhn observes two different types of periods in the development of science within a specific field: periods of “normal science” and periods of “scientific revolutions”. In periods of normal science, scientists work within a certain “paradigm” and try to show that more and more observations can be explained within this paradigm. They do not focus on falsification.

Kuhn argues that this uncritical work of scientists is important for scientific development. It has the character of puzzle solving within the paradigm. Scientists actually work hard trying to show that their paradigm “works”. But falsifying observations occur and accumulate, and sooner or later the complex and arbitrary assumptions that hold the paradigm together become too many. Then there is a period of crisis and scientific revolution and a new paradigm starts to emerge, replacing the old one.

The development of science according to Kuhn can be described in the following model:




**Main sources and suggested further reading:**
Bunge (1967), Hempel (1966), and Kuhn (1962).

**References**

Bunge, Mario. (1967). *Scientific Research*. Berlin: Springer.
Collingwood, Robin George. (1946/1994). *The Idea of History*. Oxford: Oxford University Press.
Chalmers, Alan F. (1999). *What is this thing called Science?* Buckingham: Open University Press.
Hempel, Carl. (1966). *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall.
Kuhn, Thomas S. (1962). *The Structure of Scientific Revolution*. Chicago, Ill.: University of Chicago.
Lundequist, Jerker. (1999). *Tools of Scientific Thinking*. (Forskningspublikation 1999:2). Stockholm: KTH, School of Architecture.
Nachmias, Chava & Nachmias, David. (1992). *Research Methods in the Social Sciences*. London: Edward Arnold.
Peirce, Charles Sanders. (1931—35 & 1958/1992). *The Collected Papers of Charles Sanders Peirce*. [CD—ROM]. Past Masters. InteLex Corporation 1992.
Popper, Karl R. (1972). *The Logic of Scientific Discovery*. London. Hutchinson.
Rosing, Hans. (1994). *Vetenskapens logiska grund*. Hangö: Schildts Förlag
Russell, Bertrand. (1912/1976). *The Problems of Philosophy*. Oxford: Oxford University Press.
Winch, Peter. (1958/1990). *The Idea of a Social Science and its Relation to Philosophy*. London: Routledge & Kegan Paul.
Windelband, Wilhelm. (1915). *Preludier*. Stockholm: Bonniers.
von Wright, Georg. (1971). *Explanation and Understanding*. New York: Cornell University Press.

# Research Methodology

## STRATEGY, METHODOLOGY, AND METHODS

I will use the terms “strategy”, “methodology”, and “methods” with specific meanings. Methods or techniques are used to collect and analyse data. Participant observation, questionnaires, interviews and archival records are examples of data collection methods. A methodology is a recommended set of methods for collecting and analysing data. It also has a standard for the validation of findings. A strategy, finally, links methodology to theory.

Many authors use other conceptual frameworks than I do. Martyn Denscombe (2003), for instance, does not make a distinction between strategy and methodology.

### Research strategies

Reality is complex and unlimited. Some kind of strategy is needed to focus a scientific investigation. I identify three principal strategies for reducing data: a reduction which is necessary to make the empirical world amenable to investigation *(figure 1)*.

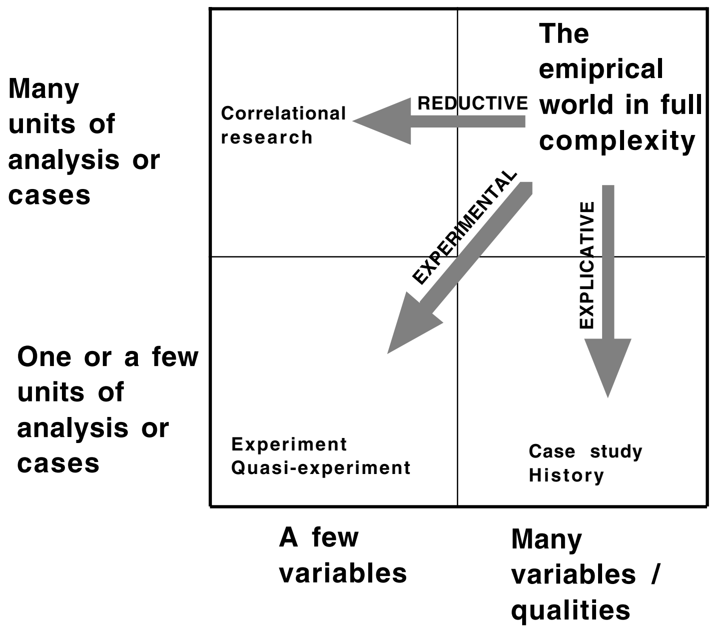



*Figure 1. Three strategies for focusing empirical research by reducing the units of analysis (cases), the number of variables (qualities), or both. The three strategies imply different methodologies. Correlational research, experiment, quasi-experiment, case study, and history are examples of research methodologies.*

One strategy is the *experimental strategy*. The word "experiment" originates from the Latin word *experiri* (to try). The aim of the strategy is to reduce the units of analysis to only one—an experiment—and the number of variables to a few. In an experiment the unit of analysis is an operation, which is carried out under controlled and restricted conditions. The scientist must be capable of manipulating, varying and measuring the relevant variables. Therefore, the experiment should preferably be conducted in a laboratory or in circumstances where the influence of external variables can be either eliminated or calculated. Theories (hypotheses) about general causal relations in nature are tested in experiments with a limited number of controlled variables. The experimental strategy has dominated scientific research since the $16^{th}$ century.

The experimental strategy is very useful for explaining general laws governing processes in nature, but cannot generally be used in the context of historical processes. The experimental strategy presupposes that every experiment may be repeated under identical conditions, but historical processes are characterised by qualitative change. Therefore historical processes are obstacles to the principle of experimental replication.

The second strategy is the *reductive strategy*. "Reductive" derivates from the Latin word *reducere* (diminish, confine). The aim of the strategy is to use many units of analysis, but reduce the number of variables. The variables analysed are of a higher order, which means that they are defined out of context: they are not related to a specific situation. This strategy makes it possible to formulate context-independent theories about patterns and surface structure.

The reductive strategy became a viable alternative when the science of mathematical probability and its application—statistics—developed. In the sciences the reductive strategy is applied, for instance, in a statistical investigation or in a survey. Prediction from statistics is based on induction. It thus has a limitation: a principle that actually *is* active may be identified, but *why* it is active will not be revealed.

This might be illustrated by an example. Imagine that a man jumps from the top floor of a skyscraper. He observes that he passes the $42^{nd}$ floor without any problems. He then observes the same passing the $41^{st}$, the $40^{th}$, and so on. When he has passed as many as 90% of the floors, he predicts that he will not have any problems at all passing all the floors of the building. This prediction is false.

The *explicative strategy* is a third strategy. This is practised within the humanities and social sciences. "Explicative" originates from *explicare* (to unfold). This strategy is effective in order to understand complex phenomena in the context in which they occur. The aim of the explicative strategy is to focus our attention on one unit of analysis—a case—simultaneously accommodating as many relevant variables and qualities as possible. Thus the case may be understood as a complex whole.

The experimental and the explicative strategies are similar insofar as they are both limited to one unit of analysis. They also both apply a procedure for the purposeful selection of the unit of analysis. In an experiment, the unit of analysis is a specific and controlled *event*, and in a case study, the unit of analysis is a case. Generalisations from an experiment or a case study have to be analytical; that is, from an experiment to a general law, or from a case study to either a theory (concepts), a specific category of cases, or, by comparison, to other cases. The reductive strategy applies a representational sampling procedure, where the sample represents a whole population, and therefore generalisations are statistical. Generalisations are made from a representational sample to a whole population.

The experimental strategy is different from both the reductive and the explicative strategies in one respect: the experiment is preferably performed in a laboratory setting, where the influence of external factors may effectively be eliminated. When the reductive or explicative strategy is used, data are collected from real life. The difference between these strategies is that in the case of the explicative strategy, data is preserved within its context, while in the case of the reductive strategy, data is separated from its context.

**Main sources and suggested further reading:**
Johansson (2002).

## ONE UNIT OF ANALYSIS

I will now discuss methodologies that focus on one unit of analysis: the experimental and the explicative strategies.

### Experimental and quasi-experimental testing of hypotheses

If we want to test a hypothesis (**H**) we must first deduce an implication from the hypothesis. This has to be a prediction of some effects that should be easy to observe and should occur if the hypothesis is true. The observable effects are called test implications from the hypothesis. The test implications (**I**) are normally of a conditional character. They tell us that under specific testing conditions (**C**) a certain effect (**E**) will occur: if the conditions **C** are present, then the effect **E** will be observable.

We can illustrate this with an example from Semmelweis:
**H**: some substance that is transferred through contact from dead bodies to patients causes childbed fever.
**I**: if everybody washes their hands with disinfectant before they have any contact with patients (**C**), then the death rate in childbed fever will decrease (**E**).

It must be possible to practically realise the conditions. In this example, one factor has to be controlled in order to conduct the experiment: washing hands with disinfectant. When it is possible to realise the conditions and to observe if the expected effects occurs, then experimental testing is possible.

If a hypothesis proves to be false, the experimenter must follow the pattern of reasoning illustrated here:

If **H** is true, then **I** is true
But **I** is (as is shown by facts) false

---

**H** is false

If a hypothesis proves to be false, the next step is to come up with a new hypothesis and test it, until a hypothesis proves to be true.

If **H** is true, then **I** is true
**I** is true (as is shown by facts)

---

**H** is true

This procedure is called the hypothetic–deductive method.

An experiment or a quasi-experiment aims at demonstrating causal relationships. The demonstration of causality involves three distinct operations (Nachmias & Nachmias 1992). The first operation demonstrates *covariation*: it is necessary to show that two or more phenomena vary together. The second operation demonstrates that the observed covariation is *non-spurious:* a covariaton is non-spurious when the relation between two variables cannot be explained by a third variable. Suppose we observe that there is a relation between the number of firefighters at a fire site and the amount of fire damage: the more firefighters at the fire site, the greater the amount of fire damage. Since the firefighters did not cause fire damage, the amount of fire damage cannot be explained by the number of fire fighters. There is a covariation between the number of firefighters and the amount of fire damage, but it is a spurious relation. There is a third factor, the size of the fire, that causes both the high number of firefighters and the high degree of fire damage. The third operation is time order. This requires the researcher to demonstrate that the assumed cause occurs first or changes prior to the assumed effect.

In the discussion of experimental and quasi-experimental research designs I use a few abbreviations:

**X** = the exposure of a group or individual (in the social sciences), or a substance, particle or such entity (in the natural sciences), to an experimental variable or event, the effects of which are to be measured.
**O** = some process of observation or measurement.
**R** = random assignment.

There are three experimental designs called *true experimental designs*. They are *the posttest-only control group design*, *the pretest-posttest control group design,* and *the Solomon four-group design*. In a true experimental design, two groups are selected by random assignment. In one of the groups—the experimental group—**X** is introduced. The other group—the control group—is not exposed to **X**. The random assignment ensures that the groups are equal in all other respects than the exposure to **X**.

*1. The posttest-only control group design.*

| | | | |
|---|---|---|---|
| **R** | **X** | $\mathbf{O_1}$ | Experimental group |
| **R** | | $\mathbf{O_2}$ | Control group |

In the posttest-only control group design, the experimental group is exposed to **X**, and observations are subsequently made of both groups. The effect of **X** is reflected in the difference between $\mathbf{O_1}$ and $\mathbf{O_2}$.

*2. The pretest-posttest control group design.*

| | | | | |
|---|---|---|---|---|
| **R** | $\mathbf{O_1}$ | **X** | $\mathbf{O_2}$ | Experimental group |
| **R** | $\mathbf{O_3}$ | | $\mathbf{O_4}$ | Control group |

In the pretest-posttest control group design, observations are also made before **X** is introduced. The observations made before the occurrence of **X** will indicate whether there are differences between the groups at that time. The effect of **X** is reflected in the difference between $\mathbf{O_2}$ and $\mathbf{O_4}$. This design offers a better control of some external variables: if there are external variables that might have caused an effect that is observed in the difference between $\mathbf{O_2}$ and $\mathbf{O_4}$, they would also produce a difference between $\mathbf{O_3}$ and $\mathbf{O_4}$.

*3. The Solomon four-group design.*

| | | | | |
|---|---|---|---|---|
| **R** | $\mathbf{O_1}$ | **X** | $\mathbf{O_2}$ | Experimental group 1 |
| **R** | $\mathbf{O_3}$ | | $\mathbf{O_4}$ | Control group 1 |
| **R** | | **X** | $\mathbf{O_5}$ | Experimental group 2 |
| **R** | | | $\mathbf{O_6}$ | Control group 2 |

The third true experimental design combines the two first designs. In this design the effect of **X** is replicated in four different ways:

$\mathbf{O_2}$—$\mathbf{O_1}$
$\mathbf{O_2}$—$\mathbf{O_4}$
$\mathbf{O_5}$—$\mathbf{O_6}$
$\mathbf{O_5}$—$\mathbf{O_4}$

***Pre–experimental designs***

For moral or practical reasons in the social sciences it is often not possible to randomly assign persons to experimental groups. On many occasions people must be asked for their permission before they are exposed to **X,** and this means that the assignments to the groups are no longer random. These experimental designs are called *pre-experimental designs*.

*4. The static group comparison.*

| | | | | |
|---|---|---|---|---|
| | | **X** | $\mathbf{O_1}$ | Experimental group |
| | | | $\mathbf{O_2}$ | Control group |

The *static group comparison* design is the same as the *posttest-only control group* design, except that the assignments to the groups are not random. Since the selection factor is not controlled, it may be biased. If there is a difference between $\mathbf{O_1}$ and $\mathbf{O_2}$, this may be explained by the different recruitment of persons making up the groups may be the explanation:. The groups might have differed in any case, without the occurrence of **X**.

*5. One-shot case study.*

| | | |
|---|---|---|
| **X** | $\mathbf{O_1}$ | Experimental group |

In the one-shot case study there is no control group and an observation is made after the introduction of **X**. This design is useless in demonstrating causal relationships, and since that is the aim of an experimental design, this design is not recommended.

However, in a case study the researcher has to deal with a situation of this kind. Observations are made when things have happened or are happening. The difference between a one-shot case study in experimental design and a case study in a naturalistic setting is that the case study researcher does not focus the study on one isolated variable but on the case as a complex whole within its context. Therefore, the case study researcher has to use another methodology: case study methodology.

*6. One group pretest-posttest design.*

| | | | |
|---|---|---|---|
| $\mathbf{O_1}$ | **X** | $\mathbf{O_2}$ | Experimental group |

When an observation is made, not only after **X**, but also prior to **X**, it is possible to observe a difference between $\mathbf{O_1}$ and $\mathbf{O_2}$ that might be caused by **X**. However, there is a problem, which is that the effect might also be caused by other external factors, which are out of control.

***Quasi–experimental design***

Sometimes in the context of the naturalistic social sciences, the researcher does not have full control over the scheduling of experimental stimuli, X, which is a condition of a true experiment. Such situations are regarded as quasi-experimental designs. A quasi-experimental design requires improved data collection procedures through a periodic measurement process—a time series measurement.

The essence of time series design is that an experimental change, which might be beyond the full control of the researcher, is introduced into a periodic measurement process. The effect of X is then indicated by a discontinuity in the measurements recorded in a time series.

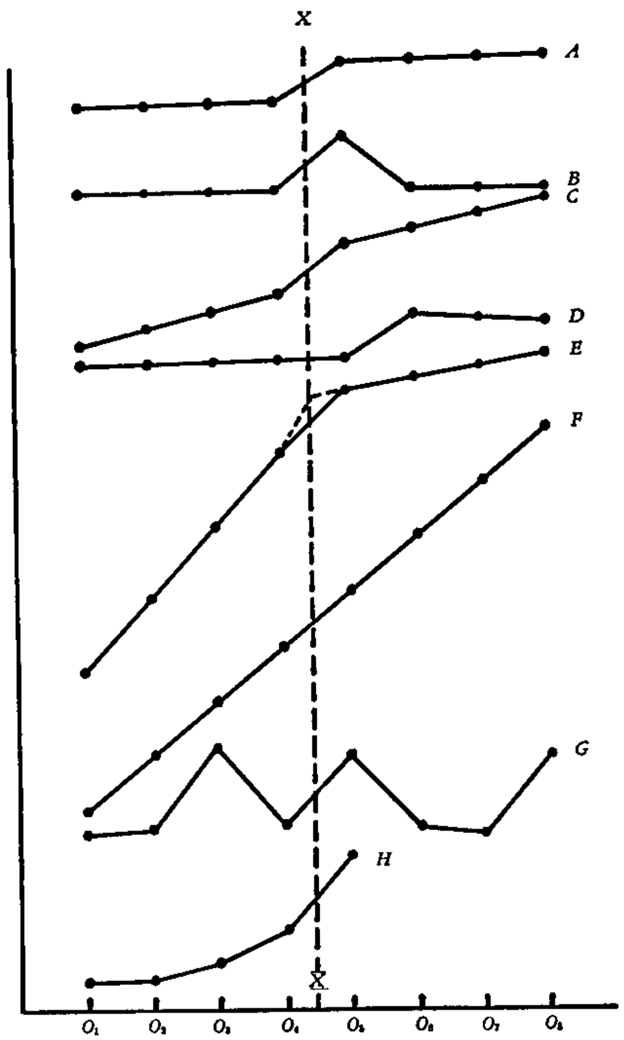



*Figure 2. This figure illustrates a possible outcome pattern from the introduction of an experimental variable at point X into a time series of measurements ($O_1$—$O_8$). From Campbell & Stanley 1966.*

Figure two shows that, except for **D**, the difference between $\mathbf{O_4}$ and $\mathbf{O_5}$ is the same for all time series. Still, the legitimacy of inferring an effect varies a lot. In **A** and **B** there is an obvious effect of **X**. But in **B**, the effect is reverting. In **C** the effect is less obvious: it is a temporary accelerated change. In **D** there may be a delayed effect of **X**. In **E** there is a change of tendency that seems to be caused by **X**. In **F**, **G,** and **H** inferring an effect of **X** is completely unjustified.

**Main sources and suggested further reading:**
Campbell & Stanley (1966), Nachmias & Nachmias (1992).

## Case study methodology

The case study appears within a variety of research traditions—such as sociology, ethnography, and biography—and within spheres of practical activity. Journalism, social work, healthcare, law, economics, architecture, planning and environmental studies are examples of areas where the case study is applied abundantly.

A case study is defined by the choice of object to be studied—a functioning entity in its naturalistic setting—and there are correspondingly different case study methodologies.

Before we continue this discussion we need some way of categorising case studies which have differing characteristics. I choose a division into three categories. This respects the logical ground of each methodology. The categories are deductive, inductive and abductive case study methodologies. These divisions correspond to the three different ways of deriving conclusions or convictions: deductively, inductively and abductively. With deduction, we apply a set of hypothetical conditions to an imagined case in order to deduce a certain fact. If the case is then implemented, the theoretically derived fact can be compared with an observed fact to determine whether the imagined case—in the form of a hypothesis about a causal relationship—is correct. Inductions are derived from a number of facts, and are attempts to find patterns that imply a rule. Abductions, finally, are derived from unexpected facts, and posit a theory that *may* be correct in constructing a *possible* case—a case that makes the unexpected fact appear intelligible. These different forms of reasoning deliver different types of answers: a derived *fact*, to be compared with an observed fact; an applicable *rule*; and a possible *case*. They provide answers to different kinds of questions: Does the rule apply? What rule is applicable? And what is the case?

### *Deductive case study methodology*

The deductive case study methodology has, as an ideal, the scientific experiment. Thus, it also has logical positivism as its basic paradigm: it assumes the existence of an objective reality, independent of us, that we can investigate. As with an experiment, the starting-point is a hypothesis about what rules apply in a particular case. Or, rather, there may be two competing hypotheses that are mutually exclusive: that is to say, if one is true, the other is necessarily false. The course of action is then to seek evidence that verifies or falsifies the expected consequences of the hypotheses. The hypotheses lead the investigations: they indicate what facts are relevant. Facts are validated through triangulation.[1] As a large number of different data collection methods are used, facts are validated primarily through method triangulation: both quantitative and qualitative. The

---

[1] The concept of triangulation will be explained later.

analysis methods often derive from quasi-experiments, for instance, time–series analysis and pattern–matching, as shown in f*igure 3*.

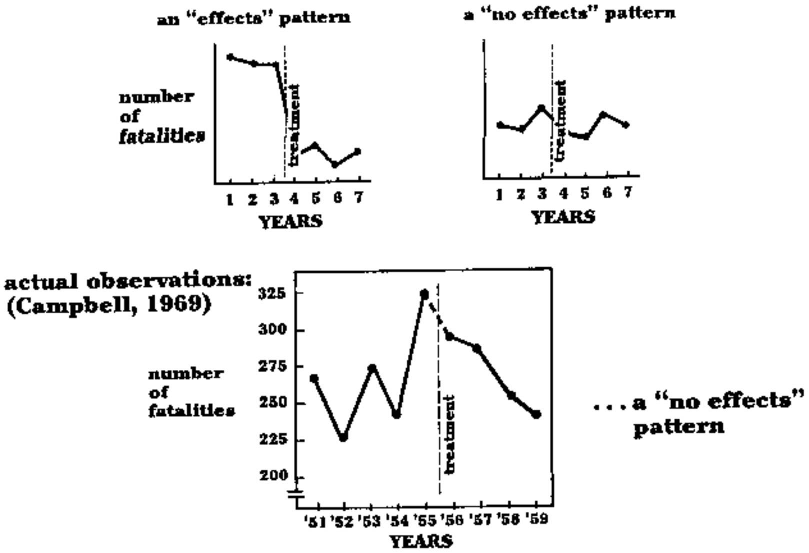



*Figure 3. An example of pattern-matching, from Yin (1994).*

The deductive case study methodology is best explained by Robert Yin in his book *Case Study Research*, which first came out in 1984. Yin is himself schooled within experimental psychology. His mentor, Donald Campbell, who wrote the preface to the methodology book, is a prominent authority in the areas of experimental and quasi-experimental research design.

***Inductive case study methodology***

Inductive case study methodology appears predominantly within ethnographic research traditions. In the ethnographic case study, the case consists of a social group or societal formation, and questions that are posed concern cultural factors. Fieldwork in the form of observation through direct participation is the preferred data collection method. This often becomes the dominant method. The sole dependence on observation serves to make the case studies only slightly triangulating with respect to method, but this can be compensated for by collecting data over a long period of time—prolonged fieldwork—and data triangulation is then achieved. Reality is usually perceived as socially constructed.

Within sociology the case study originates from accounts of cases of social work ("case history" or "case work"). The sociological case study was developed early within the Chicago School (University of Chicago). The qualitative methodology of the Chicago School has its roots in the early twentieth-century and was influenced by American pragmatism. It enjoyed a heyday from the 1920s to the 1940s. Field studies in the form of participant observation were carried out in the immediate surroundings of the university campus, where city life was investigated and explored with anthropological methods, as if it were an exotic environment.

During the 1940s, an animated debate broke out between Herbert Blumer of the Chicago School and statisticians who argued in favour of quantitative methods. "Proponents of case study argued that it yielded 'understanding', whereas proponents of statistics claimed that was the only way to science" (Whyte 1943/1993). One is inclined to say that the statisticians won the debate. Within all the social sciences, the reductive strategy speedily gained ground relative to the explicative strategy. The social sciences were to become dominated by quantitative methods.

Within the tradition of ethnographic research, the case study developed in two directions after this positivistic criticism. One strand developed into action research, which to an even greater extent aims to construct a social reality. The other complies with the criticism and aims at a more scientific systematisation of the case study methodology. The result was a new methodology: *Grounded Theory*.

*Grounded Theory* differs from other forms of ethnographic methodology, as, like deductive methodology, it assumes an objective reality ready to explore. Ethnographic methodology generates a description and interpretation of a case. *Grounded Theory*, on the other hand, is more focused on systematically building up a theory about the case. Grounded Theory also differs in its emphasis on method triangulation as a means to collect data—above all, qualitative data—and recommends interviews, observations, and studies of documents as suitable methods for this. The methodology was first introduced in *The Discovery of Grounded Theory* by Glaser & Strauss (1967), and is characterised by rigorous procedures for data coding and theory construction. The two authors represent entirely different schools of research. Barney Glaser's background was at Columbia University where, in the 1950s, a model for *quantitative* inductive theory generation was developed, while Anslem Strauss was a student of Herbert Blumer and was schooled in *qualitative* research, largely based on fieldwork, at the University of Chicago.

A case study in the ethnographic research tradition, based on inductive methodology, assumes a case (normally, a group of culturally and geographically coherent people) and collects large amounts of data about the case through fieldwork. Based on the collected data, a description and an interpretation of, or a theory about, the case is then inductively constructed. Unlike the deductive model, there is no hypothesis that directs the data collection. This is particularly stressed in *Grounded Theory*. The theory is the result, and this may consist in making concepts more precise.

### ***Abductive case study methodology***

The abductive case study is similar to the inductive model. A main difference, however, is that, in the inductive case study, we begin with a conception of what the case is, collect facts, and induce an interpretation or theory, whereas the abductive case study emanates from (meagre) unexpected facts, and tests experience-based or innovative models in order to construct a plausible case. The problem is analogous to a riddle. The riddle springs from a limited access to facts, and one has to solve the riddle to make sense of it. Whereas inductive methodology depends on abundant facts, the abductive methodology is based on incomplete facts.

Case studies within the domain of history are often abductive. One has to make sense of the past using accessible (and often questionable) facts. All historical studies, however, are *not* case studies. This is because they sometimes lack a sufficiently specific case to serve as the objective of a study. Historical research was, like the social sciences, inclined to quantitative and "neo scientific" methods during the 1950s and 1960s. Biographical studies and monographs on buildings or places are, however, obviously studies of unique cases. They are historical case studies.

The journalists Bernstein and Woodward's case study (1974) about the effects of the burglary in the democrats' headquarters in Watergate is largely abductive in its nature. One Saturday morning in 1972, five burglars were caught red-handed. They looked like businessmen, wore surgeons' gloves, and carried walkie-talkies, cameras, picklocks, and bugging equipment. They began their operation with a lobster meal at a luxury restaurant on the ground floor of the Watergate complex. All five were ex- ¾ agents. One of them, James McCord, was responsible for the security in the re-election campaign of the then president Nixon. Indisputably unexpected facts. "The two reporters stood in the middle of the newsroom and looked at each other. What the hell do you think it means? Woodward asked. Bernstein didn't know". This is where the riddle was formulated that drove the journalists: What does this mean? What is the actual case? It appeared that the case was that Nixon's closest men, with Nixon's knowledge, had, for a long time, organised and financed systematic smearing of the democrats' candidates, and sabotage of their campaign work. The Watergate burglary was only one trick of many. Nixon was forced to resign (*figure 4*).

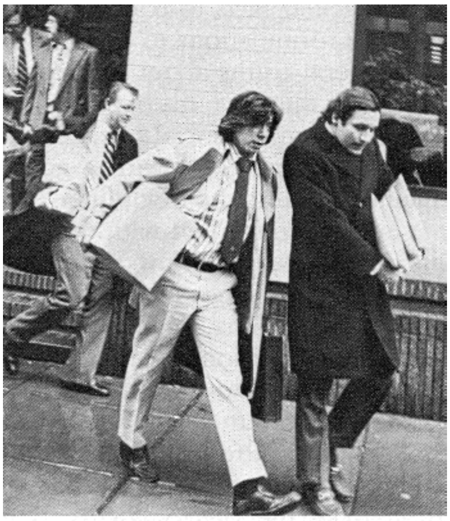


*Figure 4. Carl Bernstein and Bob Woodward leave the Washington Post headquarters, carrying material for the case study that passed the sentence on Nixon (Davidson & Lytle 1982).*

**How to generalise from a case?**

In deductive case studies the procedures of the experiment are emulated. Theory—in the form of, for instance, universal law, microtheory, concepts—or a hypothesis that states a principle that should be valid in the case, is tested.

Imagine a detective called to a murder case in a flat where he discovers the victim of a murderer. There has been a fight. Based on his experiences from similar cases he formulates two hypotheses that could both explain the case, but are mutually exclusive. One (1) proposes that this is a case of revenge, and the other (2) that this is a case of robbery. He then looks for evidence that verifies or falsifies either of his hypotheses. If (1) is true, the murderer knew the victim and must be found among friends, relatives and other acquaintances of the victim. If (2) is true, the murderer probably did not know the victim. The search for evidence is guided by the hypotheses. If both are falsified a new one must be constructed.

Theory-testing in a case leads to better knowledge of the domain in which the theory is valid: we generalise from a case and to the applicability of a theory.

In an inductive case study, theory is generated from data in the case. Normally, theory should be understood as concepts or systems of concepts and micro level theory. In this case, generalisations are again made from a case to theory, not through testing, but through generation.

A case study may be intrinsic: not aiming at generalisations, but at understanding—or learning from—the specific case. In such cases, generalisations are left to be made by the reader of the case study. By "naturalistic" generalisation a practitioner—a planner, for instance, facing an actual planning problem—may generalise from a case to another case by comparing the case study with his/her own case. "Naturalistic generalisations are conclusions arrived at through personal engagement in life's affairs or by vicarious experience so well constructed that the person feels as if it happened to themselves." (Stake 1995).

The processes of testing theory, generating theory and (re)constructing a case might very well be present in the same case study.

**Main sources and suggested further reading:**
Johansson (2002), Yin (1984/1994), Stake (1995).

**References**

Bernstein, Carl, & Woodward, Bob. (1974). *All the President's Men*. New York: Simon & Schuster.

Campbell, Donald T., & Stanley, Julian C. (1966). *Experimental and Quasi-experimental Designs for Research*. Boston: Houghton Mifflin Company.

Collingwood, Robin George. (1939/1970). *An autobiography*. Oxford: Claredon Press.

Glaser, Barney G., & Strauss, Anshelm L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Pub. Company.

Johansson, Rolf. (2002). Ett explikativt angreppssätt – Fallstudiemetodikens utveckling, logiska grund och betydelse i arkitekturforskningen. *Nordic Journal of Architectural Research, 15,* (2), 19–28.

Nachmias & Nachmias. (1992). *Research methods in the Social Sciences*. London, Melbourne and Auckland: Edward Arnold.

Stake, Robert E. (1995). *The Art of Case Study Research*. Thousand Oaks: Sage.

Yin, Robert K. (1984/1994). *Case Study Research: Design and Methods*. Thousand Oaks: Sage.

Whyte, William Foote. (1943/1993). *Street Corner Society: The Social Structure of an Italian Slum* (4th ed.). Chicago: University of Chicago Press.

## CASE STUDY METHODOLOGY

### What is a case study?

A case study is both a process of learning about a case and the product of our learning from a case: a case study report. Case studies are common in professions which involve practical activities such as social work, health care, economics, law, management, and education. Architecture, planning, and environmental engineering are also such practical professions.

The case study, as we know it today, has been developed within the social sciences. There are two traditions which have influenced case study methodology: one is the tradition of qualitatively oriented research from the humanities, a second is the quantitatively oriented experimental tradition from the natural sciences. Case study methodology often aims at merging the two traditions. Michael Quinn Patton (1990), for instance, claims that

> Rather than believing that one must choose to align with one paradigm or the other, I advocate a paradigm of choices. A paradigm of choices rejects methodological orthodoxy in favor of *methodological appropriateness*.

In a case study, qualitative methods are normally used in combination with quantitative methods.

Robert Yin (1994), a case study researcher in the experimental tradition, defines case study research in the following way:

> *A case study is an empirical enquiry that:*
> • investigates a contemporary phenomenon within its real-life context, especially when
> • the boundaries between the phenomenon and the context are not clearly evident.

Yin emphasises the method as the most important characteristic of a case study. A case study "investigates contemporary phenomena within its real-life context". The study object is a phenomenon that is studied in its naturalistic setting.

Robert Stake (1998) advocates another criterion for the case study: "case study is not a methodological choice, but a choice of object to be studied". And the object of study must be a case. The case, in turn, must be "a functioning specific". In Stake's definition every study that has a case as the object of study is a case study regardless of the methods used.

Yin makes a distinction between different types of case studies. A case study may be holistic or embedded. A holistic case study is one which is focused on the case as a unit of analysis. In an embedded case study, the case still functions as the unit of analysis, but there are also subunits of analysis within the case. If a housing area is the case, the households in the area may be embedded units of analysis, and the case study is thus an embedded case study.

Yin also draws a distinction between a single-case design and a multiple-case design. In a single case design, only one case is used to address the questions which are being researched; in a multiple-case design two or more cases are studied.

Robert Stake distinguishes between three types of case study: collective, intrinsic, and instrumental case studies. A collective case study is equivalent to a multiple-case study. The aim of an intrinsic case study is to learn about a specific case without generalisation. The aim of an instrumental case study is to generalise the findings.

**What is a case?**

A case is a phenomenon of some kind which emerges in a limited context. The case is thus a phenomenon in its context. The phenomenon must be complex in order to be the object of a case study. And when it is complex, it is not necessarily clear where the boundary lies between the phenomenon we wish to focus on and its context: it is normal that the boundary alters during the process of the study. Cases can have differing natures: they may be determined in social, spatial, or temporal terms.

A socially determined case may be *an individual* in a defined context: an environmental engineer in his profession, for instance. A case may also be *a role*; let us say, for example, the role of an environmentalist in a community. Or it can be *a small group*, such as the planners in a municipality. The case may be *an organisation*: an ecologist association, for example. Finally, the case may be a limited social context, such as *a village community* or *a commune/collective*. In all these examples it is the type and size of the social unit which determines the case.

Spatial determination is another basis for defining a case. A case may, for instance, be *a building*, *a garden*, *a housing area*, or *a town*. A third way to define a case is to limit it temporally. In this case, the case may be *an event*, *a process* (which has a beginning and an end), *an episode*, *a meeting*, or *a period of time*. An accidental outlet from a chemical factory may be a case.

A case is normally determined by a combination of social, spatial, and temporal aspects. The case is that which the case study eventually intends to say something about. It is the complex phenomenon in its natural context, our understanding of which the case study contributes to.

## How to select a case to study

In an "intrinsic case study" the case is often given: a case exists which it is of sufficient interest to study for its own sake. A criminal investigator has his or her case given in that it is determined by the fact of the crime that has been committed; a social worker often encounters particularly difficult cases which require more in-depth study; and in evaluating a project that we are responsible for, the case is also given. The choice of case is not a methodological question in such types of case study: it is simply necessary to understand the case, where this forms the ground for action in a particular situation.

In an instrumental case study, however, the choice of a case is of the greatest significance, and is crucial in determining the possibilities of generalising to theories or of generalising over types of cases. The selection of a case does not give us a representative sample which should represent a larger population about which one can potentially generalise: it is, rather, a choice based on considerations about, or a theory concerning, why a particular case would be particularly instructive or important to study. The case is selected for particular purposes. Within different research traditions, the method of choosing a case is usually called analytic, strategic, or theoretical. The motive for choosing a case may, for example, be that the case is critical in either challenging or confirming a theory. Or it may be that the case is unique as it only occurs as a single case, or typical in that it represents a category of cases, or it might be that the case is sufficiently rich in information that there is reason to believe that there is much to learn from it. With some case studies, there are a large number of additional motivating factors for choosing the case. It is always important to make explicit the arguments for selecting a specific case to study.

The selected case must be information-rich and selected for this purpose. Patton (1990) gives a number of examples of purposeful sampling, the most common ones being:

- *Extreme or deviant case sampling.* The purpose of such a study is to learn from highly unusual manifestations of the phenomenon which is of interest, such as outstanding successes or notable failures.
- *Typical case sampling.* The purpose of such a study is to illustrate or highlight the typical, normal, or average.
- *Critical case sampling.* The purpose of such a study is to verify or falsify a theory (hypothesis) through an application of the logic which is used in experiments.
- *Snowball or chain sampling.* This is often used within a case to select persons to interview. Each interviewee is asked to suggest who to interview next to gain more information.
- *Criterion sampling.* Cases that meet some criterion are selected.
- *Combination or mixed purposeful sampling.* Two or more principles of purposeful sampling are used in combination. This provides a kind of triangulation ("triangulation" will be discussed later in this lecture).

## How to collect evidence

The most common data collection methods in a case study are as follows (Yin 1984/1994).

### *Documentation*

Documents can take many forms. There are a variety of documents that might be collected and analysed in a case study:

- Letters, memoranda, and other communication.
- Agendas, announcements, minutes of meetings, and other written reports of events.
- Administrative documents: proposals, progress reports, and other internal documents.
- Formal studies or evaluations of the same “site” under study.
- Newspaper clippings and other articles appearing in the mass media.

Documentation has the advantage of being:

- Stable: it does not change and can be reviewed many times.
- Unobtrusive: documentation is not created as a result of the case study.
- Exact: names, places, times, and other details are often exact.
- Broad in coverage: in a series of agendas from meetings, for instance, a long time span is covered.

Weaknesses are:

- Retrievability: documentation might be difficult to retrieve.
- Bias: if collection is incomplete the documentation will be biased through the way in which it has been selected, and there may be a reporting bias that reflects bias of author.
- Access: access may be deliberately blocked due to, for instance, security or privacy reasons.

### *Archival records*

The most common type of archival records consists of computerised quantitative information such as statistics or survey data.

Archival records have the same advantages as documentation, with the addition of being precise and quantitative. Their weaknesses are the same as of documentation.

***Interviews***
An interview may be more or less structured.

**Unstructured**

- Listening to other people's conversation: a kind of verbal observation.
- Using "natural" conversation to ask questions.
- "Open-ended" interviews: just a few open questions, e.g. "elite interviewing" (interview with someone in a position of authority).
- Semi-structured interviews: open and closed questions.
- Recording schedules: in effect, verbally administered questionnaires.
- Semi-structured questionnaires: multiple choice and open questions.
- Structured questionnaires: simple, specific, closed questions.

**Structured**

*Figure 1. The verbal data dimension according to Gillham (2000).*

In case studies, the most commonly used interviewing technique is the semi-structured interview.

The strengths of the interview as a data collection method are that it is:
- Targeted: the interview focuses on the case study topic.
- Insightful: it gives information from personal experiences.

Weaknesses of the interview are that they may be biased because of poorly constructed questions or response bias. The interviewee may be inaccurate due to poor recall, or give the answers that the interviewer wants to hear.

***Direct observation***
This is watching from "outside" in a carefully timed and specified way.

Direct observation enables us to observe events in real time and to observe the context of events. A weakness of direct observation is that the method is time-consuming and is therefore costly. Other weaknesses are the risks of selectivity and reflexivity: events may be effected by observation.

### ***Participant observation***

Participant observation shares the advantages of direct observation, but may also provide insight into interpersonal behaviour and motives. Its weaknesses are the same as those of direct observation, and in addition, there is a risk that the investigator might manipulate events.

### ***Physical artefacts***

Physical artefacts provide insight into features of a culture and technical operations. The drawback of using physical artefacts are problems concerning selectivity and availability.

### **How to validate evidence**

The most important way of making the results of a case study valid is through *triangulation*. "Triangulation has been generally considered a process of using multiple perceptions to clarify meaning by identifying different ways the phenomenon is being seen" (Stake 1998). This is so regardless of whether the study is done because of an interest in exploring the case for its own sake or whether it is instrumentally oriented. Triangulation means that one applies several—three or more—alternative approaches to each item. In connection with qualitative methods one can, according to Patton (1987), use:

(i) Data triangulation: several sources from which to collect data about the same phenomenon.
(ii) Researcher triangulation: several researchers study the same phenomenon.
(iii) Theory triangulation: the same data is analysed using different principles.
(iv) Method triangulation: several methods are used to gather data about the same phenomenon.

Method triangulation is almost always present in a case study. Different methods are used to collect data. Sometimes method triangulation is combined with other modes of triangulation.

One of the ways to increase validity through triangulation is to conduct several types of analyses (theoretical triangulation). Yin (1994) recommends four analytic methods:

(i) comparing analyses or patterns (Pattern–Matching),
(ii) shaping a story (Explanation–Building),
(iii) analysing time-series (Time–Series Analysis), and
(iv) analysing a programme's logical model (Program Logic Models).

Comparative analysis entails identifying observable patterns and comparing them with patterns which have been predicted or are known from other cases. To analyse a building by comparing it to the ideal type of that kind of building is a form of pattern-matching. To shape a story entails either building an explanatory tale which reveals the causal connections (this is what Yin is primarily alluding to) or (more in line with Stake) to shape a coherent and plausible holistic picture of a phenomenon.

Time-series analysis entails clarifying the order in which different events or activities occur, the degree to which they vary over time, or the extent to which they are extended in time. To analyse the logical model of a programme involves analysing the assumptions about connections, usually between inputs and effects, which a programme more or less explicitly contains.

Yin (1984/1994) has summarised tactics to secure validity and reliability in a case study in the following way:

| **Tests** | **Case study tactic** | **Phase of research in which tactic occurs** |
|---|---|---|
| Construct validity (establishing correct operational measures for the concepts being studied) | • Use multiple sources of evidence (method triangulation)<br>• Establish chain of evidence<br>• Have key informants review draft case study report | Data collection<br><br>Data collection<br>Composition |
| Internal validity (establishing a causal relationship) | • do pattern-matching<br>• do explanation-building<br>• do time-series analysis | Data analysis<br>Data analysis<br>Data analysis |
| External validity (Establish the domain to which a study's findings can be generalized) | • use replication logic in multiple-case studies | Research design |
| Reliability (demonstrating that the operations—such as the data collection procedures—can be repeated, with the same results) | • use case study protocol<br>• develop case study data base | Data collection<br>Data collection |

*Figure 2. Case Study Tactics for four design tests. (From Yin 1984/1994).*

### How to generalise

From an instrumental case study one can, as has been mentioned previously, generalise to a theory, a category of cases, or concrete rules of action.

It is possible to generalise to a theory in two different ways:

(1) Hypothesis-testing. The result is knowledge about the validity of the hypothesis or theory. The domain where the theory is valid will be more precisely defined.
(2) Theory generation. Theory is built from data within the case through conceptualisation.

There is also a possibility of generalising from an "intrinsic case study", which may be worth paying attention to in architectural research. In the framework of a practical discipline such as architecture, a case study, in the form of a story about a case, can parallel real experiences and become part of the reader's—the practitioner's—experiential knowledge. The reader can re-experience the case, and the case study will thus contribute to the socially constructed knowledge of the practitioner. Stake (1995) calls this type of generalising "naturalistic generalisation". Case studies provide examples which may become part of the repertoire of the practitioner who relived the case, analogous to the self-experienced case. Experiences from cases are "naturalistically generalised" by facilitating a preparedness to act in similar situations.

### Advice to a case study researcher

If you do case study research:

- give analytical arguments for the selection of the case, and
- use multiple methods for data collection, and also try, if possible, to triangulate in other respects.

### Main sources and recommended further reading

Yin (1984/1994), Stake (1998), Patton (1990), and Gillham (2000).

### References

Gillham, Bill. (2000). *Case Study Research Methods*. London: Continuum.

Patton, Michael Quinn. (1990). *Qualitative Evaluation and Research Methods* (2nd ed.). Newbury Park, California: Sage.

Patton, Michael Quinn. (1987). *How to Use Qualitative Methods in Evaluation*. Newbury Park, California: Sage.

Stake, Robert. (1998). *The Art of Case Study Research*. Thousand Oaks, California: Sage.

Yin, Robert (1984/1994). *Case Study Research: Design and methods* (2nd ed.). Thousand Oaks, California: Sage. (First edition published in 1984).