



8.1 Essence and Relevance of Causality

People have a fundamental need to “get to the bottom of things”, that is, to understand the reason behind, for example, the reasons for the course of the stars or the ways to live a happy life or the causes of economic growth. People are looking for explanations. Godfrey-Smith (2003, p. 194) puts it concisely: “To explain something is to describe what caused it”. It is therefore not surprising that questions of causality, the *search for causes and effects*, have for a long time occupied people, in particular scientists. There are differing views and comprehensive discussions on the nature and characterization of causality in the philosophy of science (see, for example, Godfrey-Smith 2003, pp. 194ff.).

In Chap. 7, the focus was on the testing of hypotheses. In a scientific context, the test of **relationships** between variables is of particular interest. This chapter deals with a special kind of relationship, so called **causal relations**, which have particular significance and—because of that—place particular demands on the nature of the relationships between variables. The first section deals with the essential features of **causality**, then types of causal relationships are outlined. Other parts of this chapter deal mainly with basic ideas about conducting experiments, which is the most common method for the study of causal relationships.

The philosophical literature has dealt with the question, “What is causality?” for nearly 400 years. Of course, this textbook does not try to discuss and understand this stream of literature in its entirety. Introductions and summaries are offered, amongst others, by Humphreys (2000), Mumford and Anjum (2013) and Psillos (2002). Even those who cannot or will not understand the details of this discussion will be easily able to assess the *relevance of causality* through a few examples. The following examples from different areas of business, society and science/technology show “that the concept of causality is a heuristic that helps us to think about our environment, organize our thoughts, predict future events, and even change future events” (Jaccard and Jacoby 2010, p. 140). Based on these examples (see Mumford and

Anjum 2013, p. 1), more general features of causality will be characterized in the following sections.

1. **Society:** Individual behaviors have consequences, for example, careless parenting is considered a possible cause of children's poor academic performance. If there was not a causal link, one could not speak of (co-)responsibility of the parents.
2. **Law:** Human behavior (for example, in traffic) can cause physical or material damage to other people. Without a causal relationship (behavior \rightarrow damage), there could be no evidence of guilt or claims.
3. **Technology:** In the case of accidents, technical defects, etc., one typically looks for the causes (causes of accidents, etc.), on the one hand to clarify the responsibility and to derive a claim settlement from it. On the other hand, one wants to learn from it and reduce or eliminate such risks in the future. This often requires the analysis of a *causal chain*, i.e., the individual steps between a cause and the resulting consequences or effects (see Sect. 8.2). Thus, the collapse of a bridge (in a nonprofessional's conception) could have come about through the following causal chain: steel reinforcement of the concrete bridge poorly protected against moisture \rightarrow rapid rusting of load-bearing parts \rightarrow instability of the bridge \rightarrow collapse.
4. **Medicine:** Medical research and practice looks for the corresponding causes of disease symptoms in order to develop a therapy (e.g., high blood pressure increases the risk of infarction).
5. **Economics:** Almost daily, the media report and analyze more or less well founded or speculative causes of current macroeconomic developments, for example, "Growing domestic demand causes economic recovery".
6. **Stock exchanges:** Here, too, one finds ongoing media coverage, the essential component of which are assumptions (or hypotheses) about the reasons for current price developments, for example, "Falling interest rates lead to rising stock prices".
7. **Management:** When assessing the performance of managers, one has to assume a (direct or indirect) cause-and-effect relationship between their actions and decisions, on the one hand, and the resulting effects on success, on the other hand.
8. **Marketing:** An example of (assumed) causal relationships in marketing decisions is the so-called realization of a sales promotion action (e.g., temporary price reduction). How could someone be responsible for the use of resources if he or she did not assume a causal link to a short-term increase in sales (causal chain: sales promotion \rightarrow stimulation of customers to trial purchases and brand change \rightarrow increased sales)?

Such considerations of causality have become quite natural to us. What are typical similarities of such (and, of course, other) causal relationships? Which characteristics entail causal relationships and then (logically) serve to decide in an empirical investigation whether a causal relationship exists or not? The first aspect relates to the **common variation of cause and effect**. Example 4 above shows that elevated blood pressure is associated with an increased risk of infarction, and in

Example 8 it is shown that increased sales promotion is associated with higher sales. In connection with the first feature is the possibility of **intervention** or **manipulation** of the (assumed) cause with the aim of achieving the desired (and assumed) effect. For instance, in Example 1, one might think about changing the behavior of parents through education or communication to attain better academic achievements of the children. In Example 4, the term “therapy” includes the attempt to eliminate the causes of a disease. As for Example 5, there are examples in the tax and subsidy policies of governments and the interest rate policy of central banks. However, there are causal relationships where such interventions are not possible. The third typical feature is the **temporal sequence** in the sense that the change of the (presumed) cause precedes the (presumed) effect. This may be a time interval in the range of seconds (e.g., in the case of a traffic accident caused by human error, see Example 2) or in the range of years (e.g., in the case of long-term damage to a bridge in Example 3). Fourthly, one assumes **the absence of alternative explanations** whose securing represents an essential and often complex problem in empirical research. Thus, in Example 1 poor academic performance could also be caused by teachers, in Example 3 the bridge could also have collapsed due to poor quality of the concrete and in Example 8 the sales figures could have increased because general demand has grown in the respective market. Only if one can exclude such (other) possible reasons for the observed effect, then can it be assumed that this effect is unmistakably caused by the assumed cause. Ultimately, there must be a **meaningful theoretical relationship** between cause and effect. Even if, in Example 6, one could observe a commonality of fluctuations of the outside temperature and the stock exchange market development—with a temperature increase regularly preceding a positive development of the stock prices, and no other possible causes for the price fluctuations being detected—still, hardly anyone would assume a causal relationship between temperature and the stock market. The following section intends to shed more light on these five aspects.

There is one important difference between the above eight examples. In some of the examples, the causal relationship relates to *specific cases*, while in others, more *general relationships are involved*. For instance, in the above examples—in law (2), there are typically case-related findings on guilt and responsibility, in medicine (4) diagnoses are made for individual patients, and individual evaluations are made of managers’ performance (7). On the other hand, Examples 3, 6 and 8 refer to causal relationships, which have more general validity beyond individual cases. Nancy Cartwright (2014) distinguishes between **singular** and **general causal relationships**. In the sciences that focus on the development and testing of theories (see Sect. 2.1), interest in general causal relationships is greater. However, in some sciences (for example, in the science of history) the focus on important individual cases plays a major role (e.g., “What were the causes of World War I?”). In addition, the analysis of individual cases may also be helpful in other disciplines in the early stages of research (see Sect. 4.3.3). In the present chapter, however, general relationships are at the center of interest, since the test of causal hypotheses (typically through experiments, see Sect. 8.3) is oriented towards general causal relationships.

Now for the first feature of causal relationships, the **common variation of cause and effect**. Causal relationships are most likely to appear when the cause and effect *vary together*. If, for example, one observes several times that interest rates fall and then economic growth occurs, then this indicates a corresponding (causal) relationship. Remember, *this speaks in favor* of a causal relationship, but it is *not evidence* of a causal relationship. If interest rates and economic growth remain constant, then no evidence of a relationship exists and if the growth changes with interest rates remaining constant, then this speaks against a relationship. A change in the cause leads to a change or a difference in the effect (Psillos 2002, p. 6).

How can we imagine the relationship between cause and effect? In science and technology, one often encounters deterministic relationships, i.e., the effect always occurs (under all conditions such as location, situation, time, etc.) after the occurrence of the cause—often in a precisely determinable manner; for example, at reduced temperature, the resistance of an electric cable decreases. Such types of relationships hardly exist in the social sciences (including marketing research). Here, statements about probabilities or (with sufficiently large numbers of cases), statements about (relative) frequencies or correlations are more common. Nancy Cartwright (2014, p. 312) summarizes the basic idea: “When a cause is present there should be more of the effect than if it were absent. That is the root idea of the probabilistic theory of causation”.

This way of establishing the relationship between cause and effect hardly differs from the analysis of relationships between variables discussed in the context of hypothesis testing in Chap. 7. Accordingly, to provide evidence for a causal relationship further requirements (see below) need to be met. Common variation of cause and effect is therefore a necessary, but by no means a sufficient, condition for a causal relationship. The well-known principle of **correlation \neq causality** applies. With regard to causality, however, it is possible to ascertain that there is no causal relationship in the absence (or non-significance) of a correlation (or other measures or relationships).

The second aspect, the possibility of **intervention/manipulation**, has important practical and methodological consequences. On the one hand, it involves the use of knowledge of causal relationships for *design tasks*: in the examples given at the beginning of this section, Example 3 measures for the construction of a bridge, Example 4 for the determination of a therapy, Example 5 for an economic policy intervention and Example 8 for the realization of a promotional activity. Causal relationships are thus in a sense “recipes”: If one understands a causal relationship, then one can shape causes in such a way that certain effects are achieved or prevented (Psillos 2002, p. 6). In *empirical investigations*, typically in experiments, the manipulation of independent variables and the observation of whether the dependent variables change in the expected manner are “classic” approaches (see Sect. 8.3). However, there are causal relationships in which this kind of observation and analysis is not possible. For example, while historians may ask for the causes of a particular event, they cannot test their assumptions through manipulation; the same is true for astronomers. In the social sciences, there are also some situations in which the manipulation of an independent variable is not possible (too much effort, high

risk) or is ethically unacceptable (e.g., because of psychological or physical harm to study subjects). In such cases one often tries to come to comparable results by means of so-called quasi-experiments (see Sect. 8.3.3).

There is an interesting relationship of the previous paragraph to a fundamental aspect of various philosophy of science basic positions mentioned in Sect. 3.1, which deals with the position of **realism**, on the one hand, and **constructivism**, on the other. If one does not assume (in a constructivist view) that a reality exists that is independent of the viewer's perceptions and interpretations, then it makes little sense to carry out experiments. Under this assumption, the manipulation of real phenomena could have little impact on concepts and theories that exist only in the minds of scientists and have little to do with reality.

Theodore Arabatzis (2008, p. 164) explains the conflict between the constructivist view and the experimental approach:

“According to the early and most radical version of social constructivism, the constraints of nature on the products of scientific activity are minimal. Data are selected or even constructed in a process which reflects the social interactions within the relevant scientific community. Therefore, one should not appeal to the material world to explain the generation and acceptance of scientific knowledge.”

The third characteristic of causality is the sequence of events in the form of **cause before effect**. Which one of the variables in a causal relationship is considered the “cause” and which one the “effect” has to be based on substantive considerations. Nevertheless, the answer is not always clear. For instance, a positive correlation between advertising expenditure and company profitability could either refer to the fact that advertising expenditure influences profitability or that profitability (by means of increased financial means) influences advertising expenditure. Here, the analysis of the *temporal sequence* can clarify matters. Basically, one assumes that the suspected cause occurs before the effect. If one observed in the example that first the advertising budgets increase and later profitability occurred, this speaks of a causal relationship “advertising expenses → profitability”. Although Hitchcock (2008) refers to some special cases in physics in which the chronology and the direction of causality do not coincide, in the field of social science such an altered sequence is not quite conceivable. This also applies to cases in which certain expected events (e.g., expectation of a new iPhone, price developments) are anticipated and responded to, because in such cases the reactions are not due to these (often quite vague) future events, but due to the previously existing conjectures.

The central idea of the fourth feature, **absence of alternative explanations**, is quite simple and plausible. If one suspects a specific cause of an effect and is able to exclude all other possible causes as alternative explanations, then only the suspected cause remains to explain the effect. Alternative explanations can be both substantial

and methodical. For example, reasons for a change in attitudes among consumers might be the impact of marketing communication, a change of values, or new experiences. However, the measured attitude change could also be due to a (systematic or random) measurement error. Researchers are usually not able to exclude all conceivable alternative explanations for a finding. Nevertheless, the research design should be designed in such a way that at least the most important alternative explanations (including the methodological ones) cannot play a role. In this context, keeping the influencing variables constant and using experimental and control groups plays an essential role in such study designs (see Sect. 8.3). By using experimental (with the presumed “cause”) and control groups (no effect of the presumed “cause”) and interpreting the results in the comparison of both groups, one achieves a situation where other predictors act in the same way in both groups. The difference between the group results can be attributed to the effect of the “cause”. The prerequisite for this, however, is that there are no systematic differences between the two groups, which is generally achieved by randomizing the group assignment.

One type of causal relationship in the form of the so-called **INUS condition** explicitly takes into account the possibility that multiple causes and specific conditions for an effect may exist. This may be more in line with many marketing research questions than a simple relationship of just one possible cause and effect. “INUS” is an abbreviation for **I**nsufficient–**N**ecessary–**U**nnecessary–**S**ufficient (see, for example, Bagozzi 1980, pp. 16ff., Psillos 2002, pp. 87ff.). What is meant by this (initially somewhat cryptic) name? “A cause may be an insufficient but necessary part of a condition that is itself unnecessary but sufficient for the result” (Bagozzi 1980, p. 17). Since the central idea might still not be easy to understand, here is an example of the following causal relationship: “Advertising messages change attitudes”:

- *Not necessary for the result:* Changes in attitudes can be due to other causes (e.g., consumer experiences). Hence, advertising is not necessary for changes in attitudes.
- *Insufficient part of the conditions:* Advertising messages alone do not change any attitudes (are therefore not sufficient), but it is only under the conditions that consumers are exposed to the message, that they show sufficiently high involvement, etc.
- *Sufficient for the result:* If the conditions (see above) apply, then the attitude change arises as an effect of advertising messages; advertising would be sufficient *under these conditions*.
- *Necessary part of the conditions:* If the advertising message did not exist, then under the given conditions, attitudes would not change. Hence, advertising would therefore be necessary in this context to change attitudes.

Figure 8.1 graphically illustrates the example of an INUS condition as outlined above.

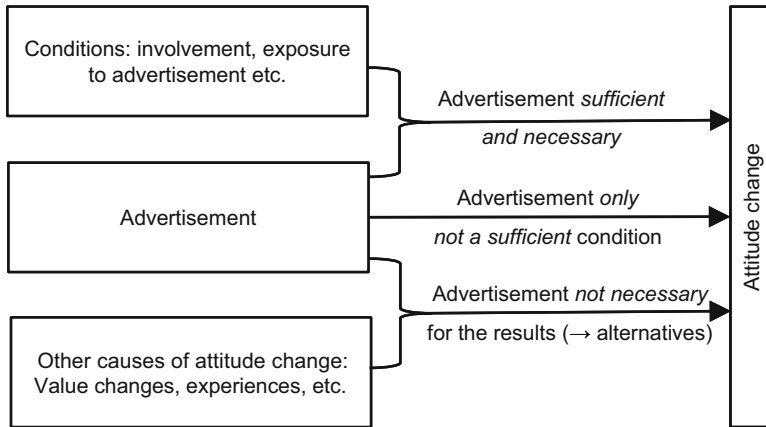


Fig. 8.1 Example of INUS conditions

Another example of Psillos (2004, p. 277) may further illustrate the somewhat complicated INUS condition:

“To say that short circuits cause house fires is to say that the short circuit is an INUS condition for house fires. It is an insufficient part because it cannot cause the fire on its own (other conditions such as oxygen, inflammable material, etc. should be present). It is, nonetheless, a nonredundant part because, without it, the rest of the conditions are not sufficient for the fire. It is just a part, and not the whole, of a sufficient condition (which includes oxygen, the presence of inflammable material, etc.), but this whole sufficient condition is not necessary, since some other cluster of conditions, for example, an arsonist with gasoline, can produce the fire.”

Let us now go back to the characteristics of causal relationships. Here is the fifth feature, where the relationship **should have a theoretical foundation**. The word “causal” already suggests that it is not about random relationships, but systematic and well-founded relationships between variables. In the social sciences, therefore, it is common to develop a chain of causation that explains and justifies the relationship between cause and effect (Cartwright 2014). For example, such a causal chain in the above described relationship between advertising and attitude change might look like this: advertising appears on TV → consumer watches and receives the message → message evokes cognitive and/or emotional responses → change of previous beliefs and evaluations → attitude change. An empirical way of analyzing such causal chains are so-called mediators, which will be discussed in Sect. 8.2.

However, with regard to the demand of a theoretical justification for a causal relationship, it should be kept in mind that this could intensify the problem of the theory-ladenness (see Sect. 3.2 and Arabatzis 2008). Corresponding empirical studies (experiments) are typically based on previously theoretically based hypotheses

and are designed accordingly. This relates to the perception and interpretation of results by the researchers, who in most cases are also “followers” of the respective theory and often try to confirm it. Peter (1991) also points out that in research practice (occasionally? often?) a research design undergoes several pretests and changes until the desired result appears, which, of course, can be problematic from an ethical research perspective (see Sect. 10.2.2).

David de Vaus (2001, p. 36) explains why a theoretical justification for the assumption of a causal relationship is essential:

“The causal assertion must make sense. We should be able to tell a story of how X affects Y if we wish to infer a causal relationship between X and Y. Even if we cannot empirically demonstrate how X affects Y we need to provide a plausible account of the connection (plausible in terms of other research, current theory etc.).”

Of the five characteristics of a causal relationship, only one—the common variation of cause and effect—directly affects the methods of statistical analysis, because it is a question of (significant) differences and changes. The last feature, the requirement of a theoretical foundation, is outside the methodological area. The three other features (manipulation, time sequence of cause before effect, and absence of alternative explanations) primarily concern the study design. “The ability to make a causal inference between two variables is a function of one’s research design, not the statistical technique used to analyze the data that are yielded by that research design” (Jaccard and Becker 2002, p. 248). Empirical methods for verifying causal relationships are typically experiments because there is close correspondence between the five outlined criteria for a causal relationship and the central elements of experimental design in experiments (see Sect. 8.3). Therefore experiments can test assumptions about causal relationships, i.e., causal hypotheses.

8.2 Types of Causal Relationships

The examination of *causal hypotheses* places particularly high demands on the methodological procedure. They lead to substantial statements in science and practice. If a researcher has determined that a particular combination of mental traits is the cause of a particular work behavior, then he or she has come a good deal closer to the goal (at least from the perspective of scientific realism) of understanding and explaining reality. When a product manager finds that certain product quality problems are the cause of decreasing market shares of a product, then he or she has found a critical starting point to solve the problem of decreasing market share.

In Fig. 8.2 there is an overview of different types of relationships between variables that either mimic causal relationships or misinterpret causal relationships.

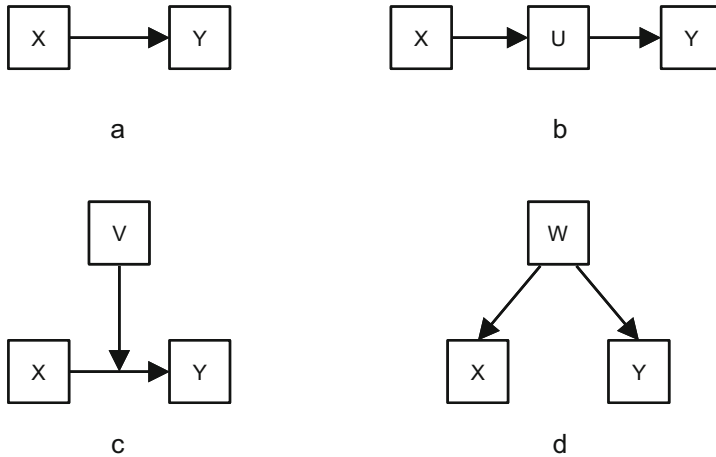


Fig. 8.2 Types of (causal) relationships between variables. (a) Direct causal relationship. (b) Indirect causal relationship. (c) Moderated causal relationship. (d) Spurious relationship (see e.g., Jaccard and Jacoby 2010)

Part a shows a simple, **direct causal relationship**, for example, the contact with advertising (cause) on the attitude to a product (effect). Part b shows an indirect causal relationship with a mediator variable (for explanation, see below). Part c shows a moderated causal relationship in which the effect of X on Y is influenced by a third variable, V (see below for explanation). Finally, part d shows a relationship that does not represent a causal relationship between X and Y because a common variation of X and Y is caused by a third variable, W. For example, the common variation of income and use of print media can be under the influence of a third variable, education. There is a danger here that the relationship between X and Y could be misinterpreted as a causal relationship.

In the moderated causal relationship, the **moderator**, a second independent variable, moderates the effect of an independent variable on a dependent variable. The influence of the independent on the dependent variable becomes stronger or weaker. The moderator can also reverse the direction of the influence: “a moderator is a qualitative (e.g., sex, race, class) or quantitative (e.g., level of reward) variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable” (Baron and Kenny 1986, p. 1174). As an example, one might think of the above relationship between exposure to advertising (X) and attitude to a brand (Y), which is moderated by the involvement with the product category: the more a consumer is involved with a product category, the stronger the effect of the exposure to advertising will be on attitudes towards a brand in that product category.

Mediators differ from moderators. Mediators designate indirect relationships between variables. Figure 8.3 shows a well-known example from advertising research (MacKenzie et al. 1986). The idea is that advertising influences attitudes towards the advertised brand. This acts, on the one hand, as a direct effect, but can

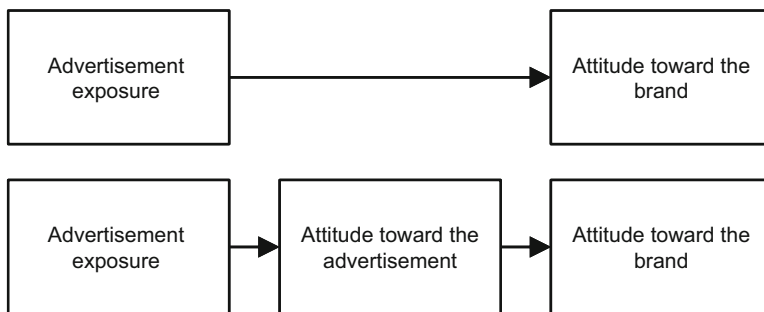


Fig. 8.3 Example of a mediator and an indirect causal relationship

also be explained by attitude to the advertisement as an indirect effect: advertising leads to the changes in attitude to the advertisement, which in turn changes the attitude to the brand. Both relationships can theoretically be justified. A direct relationship in one view (or theory) can therefore be an indirect relationship in another view (or theory).

8.3 Experimental Studies

8.3.1 Nature and Design of Experiments

Due to the five requirements for establishing causal relationships explained in Sect. 8.1, a particular study design, known as **experiment**, is commonly used. In essence, an experiment is an approach in which one or more independent variables are manipulated in such a way that the corresponding effects on a dependent variable can be observed. It is therefore a question of determining whether a certain (independent) variable is actually the reason (the cause) for a change of another (dependent) variable (effect).

Typical of experiments is the isolated consideration of the variables of interest. One does not look at a variety of factors influencing, for instance, a decision and their interactions, instead the experiment focuses only on the influence of a particular element in advertising (e.g., color or music) on the attitudes of consumers. For this reason, experimental investigations often reveal a certain *artificiality of the research design*, which is based on the exclusion of other influencing factors (→ absence of alternative explanations). Against this background, it is also easy to understand that today, one can find the results of more than one empirical study in many publications in which experiments are used. In each study, individual aspects are considered in isolation and the resulting summaries constitute a more comprehensive investigation of a topic.

Alan Chalmers (2013, p. 26) illustrates the intention of an isolated observation in the context of experiments by using the following example:

“Many kinds of processes are at work in the world around us, and they are all superimposed on, and interact with each other in complicated ways. A falling leaf is subject to gravity, air resistance and the force of winds and will also rot to some small degree as it falls. It is not possible to arrive at an understanding of these various processes by careful observation of events as they typically and naturally occur. Observation of falling leaves will not yield Galileo’s law of fall. The lesson to be learned here is rather straightforward. To acquire facts relevant for the identification and specification of the various processes at work in nature it is, in general, necessary to practically intervene to try to isolate the process under investigation and eliminate the effects of others. In short, it is necessary to do experiments.”

The major conclusions in experimental investigations can be explained by the example of a “classical” experimental design according to de Vaus (2001, pp. 48–49). The following features characterize this design:

- A pre-measure (→ sequence of cause and effect)
- Two groups: experimental group and control group (→ absence of alternative explanations)
- Random assignment of the subjects to the two groups (→ absence of alternative explanations)
- An intervention (manipulation)
- A final measurement (→ order of cause and effect)

Table 8.1 illustrates such a design. It shows the measurement times, the assignment of subjects to groups and the intervention. In both groups, attitude to a brand is pre-measured. Then, only the subjects in the experimental group are confronted with advertising for the brand. This is the **intervention** or **manipulation** of the independent variable. In the example shown, the manipulation is carried out very simply by confronting the experimental group with advertising, but not the control group. Manipulations can be diverse and can even affect mental states (such as motivations

Table 8.1 Example of a classical experimental design (according to De Vaus 2001, p. 49)

Random assignment to experimental groups	Pre measure t_1	Intervention (manipulation of independent variable) t_2	Post measure t_3
Treatment group	Attitude toward the brand t_1	<i>Exposure</i> to advertisement	Attitude toward the brand t_3
Control group	Attitude toward the brand t_1	<i>No exposure</i> to advertisement	Attitude toward the brand t_3

or emotional states). For this purpose, the different groups of subjects are influenced (or manipulated) in such a way that the corresponding mental states occur among the members of the various groups. For example, one could achieve different levels of motivation through different incentives. This process of **operationalization** (see Sect. 6.1) aims to achieve different values of independent variables. Therefore Aronson et al. (1998, p. 111) speak of “constructing the independent variable”. **Manipulation checks** usually control whether these manipulations have succeeded (e.g., whether the motivation or emotional state differs between the experimental groups). After the intervention or manipulation, the attitude to the brand is measured once more. This can occur verbally (through the use of a questionnaire) or through observations. As in the case of manipulation, one needs to consider the aspects and quality criteria of operationalization. If a significant change of attitude is measured in the experimental group only, then one would consider it as being caused by the contact with the advertisement. Are the conditions outlined above for a causal relationship given in this example?

The example fulfills the conditions for a causal relationship, if the corresponding empirical results show the expected values. This can be shown as follows:

- *Common variation of cause* (in the example “exposure to advertisement”) and *effect* (in the example “attitude to the brand” at time t_3): This condition is clearly fulfilled, since the intervention in the form of the contact with the advertisement takes place only in the experimental group. The contact with the advertisement thus varies between the groups and its measurement shows whether the dependent variable alters between both experimental groups accordingly.
- An intervention/manipulation at time t_2 is part of the experimental design.
- *Change of the cause* (in the example: exposure to the advertisement) *before change of the effect* (in the example: attitude change): This requirement is also fulfilled by the experimental design, which determines the timing of intervention and post-measure.
- *Absence of alternative explanations*: In field studies, the exclusion of all conceivable alternative explanations can hardly ever be achieved. This is certainly a weak point of experiments. Therefore, one focuses on particularly important or frequently occurring aspects of an investigation. Of central importance is the use of (comparable!) **experimental and control groups**. Ideally, these groups do not differ except for the intervention (e.g., they do not differ in terms of socio-demographic or psychological characteristics, past experience and attitudes). Therefore, different results of the final measure can only be attributed to the “cause” in the form of the intervention. In most cases, the assignment of subjects to experimental and control groups is random (**randomization**), which makes greater differences between the two groups less likely. In the example shown, the random assignment of the subjects to the experimental and control groups has (largely) excluded the fact that these groups differ systematically from one another, which could be an alternative explanation for differences in the final measure. For this reason, researchers like to work with students as subjects in experiments, because this group is largely homogeneous in terms of many

demographics (e.g., age, education, income) as well as psychographic characteristics (e.g., openness to innovation), which further reduces the risk of systematic differences. As mentioned in Sect. 6.4, however, experiments with students may be problematic if generalizability of the results is desired, but the results are systematically different from the population, for example, if students are generally more positive about advertising. Then, in the example mentioned above, students may experience an effect that may not be present in other people (non-students) or that is not so strong. Due to randomization, a pre-measurement is no longer necessary, because one can assume that the attitude to the brand at time t_1 is randomly distributed over both groups and thus on average should be approximately the same in both groups. When interpreting the results of the study, one focuses on statistically significant differences between the groups and neglects random (small) group differences with regard to the hypothesis of the investigation. Randomization as *random assignment* to experimental or control groups should be clearly differentiated from the *random selection of subjects* (*random sample*), which in experiments serve in particular to achieve external validity (see Sect. 8.3.2).

The above-mentioned alternative explanations, which are based on the methodological procedure in an experiment, are discussed in the following Sect. 8.3.2 under the heading “internal validity”. The rather complex design of experimental studies typically aims to exclude several alternative explanations (see, e.g., Shadish et al. 2002; Koschate-Fischer and Schandelmeier 2014).

- *Theoretical justification of the relationship*: The methodology cannot answer the question as to whether there is an adequate theoretical justification for an examined relationship, but a substantive consideration can. The development of an experimental design forces researchers to make deliberate considerations regarding the mode of action of independent and dependent variables (i.e., corresponding theoretical considerations). In the example used here (advertising → attitude change), the theoretical justification is established and easy to understand.

Experiments have long been widely used and are accepted methods in medicine or psychology. Accordingly, psychology-related areas of marketing research use them quite frequently (in particular, consumer research). The applications of experimental designs are typically more complex than the example given. They often examine two or three independent variables at the same time, as well as their interactions, and make manifold efforts in order to meet the requirements for the examination of causality. Please refer to the extant literature (e.g., Koschate-Fischer and Schandelmeier 2014; Shadish et al. 2002; Geuens and Pelsmacker 2017).

8.3.2 Internal and External Validity of Experiments

Chapter 6 explained the importance of the reliability and validity of a study with regard to the meaning of study results. As already mentioned, the problem is that

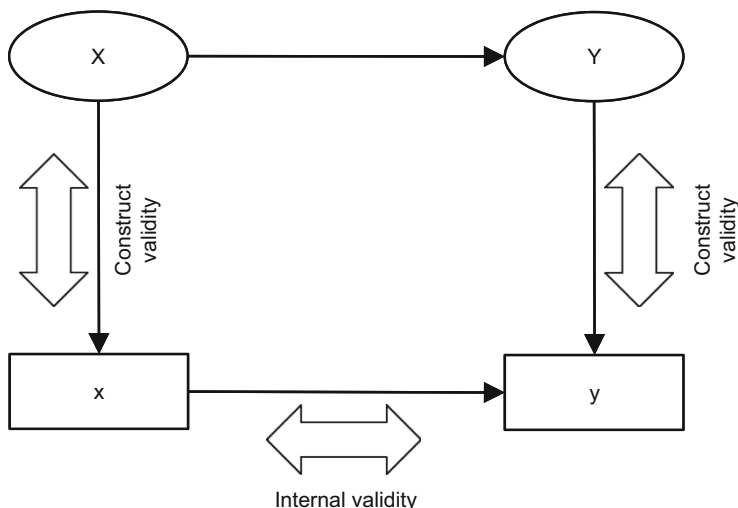


Fig. 8.4 Internal validity and construct validity in experiments

results that confirm or do not confirm a hypothesis are limited in their validity in terms of the theory tested, if these results are influenced by method errors. Concerning experiments, general considerations on the validity of studies (see Sect. 6.3) add two specific aspects: internal and external validity. The aspect of internal validity has already been implicitly addressed. **Internal validity** refers to the *elimination of alternative explanations* for the observed relationships due to the measurement process. Internal validity is thus “the validity of inferences about whether the relationship between two variables is causal” (Shadish et al. 2002, p. 508). The main question here is whether the change in a dependent variable can actually be attributed to the presumed cause, i.e., the change in an independent variable, or whether inadequacies of the methods and the measurements are responsible for the results. Figure 8.4 shows this aspect and the relation of the measured variables to the theoretically interesting concepts/constructs (\rightarrow construct validity, see Sect. 6.3.3). The lower-case letters (x, y) stand for the variables used in the study, which should be an operationalization of the corresponding concepts/constructs (upper-case letters X, Y). Construct validity is primarily related to validity in the measurement of concepts (has the concept been measured correctly?), the internal validity is concerned with the question of whether the relationship between concepts is validly represented (does the measured relationship actually exist?).

The internal validity of an experiment is mainly jeopardized by the problems mentioned below (Shadish et al. 2002, pp. 54ff.). They provide alternative explanations for the results of experiments, which are methodologically justified and that should be avoided by the design of the experimental design.

- *Selection/assignment.* The assignment to experimental and control groups might not ensure that neither group shows any systematic differences. Thus, if a

difference exists between the groups, one cannot infer the effect of the independent variables.

- *History*. Each event between pre- and post-measure may have an unwanted influence on the subjects, such as external influences that affect only a part of the subjects.
- *Maturing*. Subjects can change between two measures due to experience, fatigue etc. Therefore, it could be that subjects respond differently to stimuli over time and thus their actual effect is mitigated or nullified.
- *Change in measurement instruments*. During a study, the characteristics of the measurement instruments, including the measuring persons, may change. For example, the measurements may be made more accurate by increasing the experience of the measuring persons, or less accurate by increasing boredom during the course of the experiment.
- *Regression to the mean*. This statistical artifact can be superimposed on effects, for example, by selecting subjects with particularly extreme values, who then show (as a statistical necessity), on subsequent measures, quite “moderate” values.
- *Drop out*. Subjects may drop out during the study due to the study requirements. The affected groups are then smaller in a second measurement, which in turn can influence the result in case of a non-random drop-out.

In addition, the question arises to what extent the results of a study can be generalized. What explanatory power, for example, does a study that was carried out on German product managers have for product managers in general? What do the results of a consumer behavior experiment with 100 American students say about consumers in general? Such questions apply to the external validity of experiments. **External validity** refers to the generalizability (see also Chap. 6) of results about different persons, situations, contexts etc. External validity is therefore: “the validity of inferences about whether the causal relationship holds over variations in persons, settings, treatment variables, and measurement variables” (Shadish et al. 2002, p. 507).

Campbell and Stanley (1963, p. 5) formulate the central points of internal and external validity as follows:

“Fundamental (...) is a distinction between internal validity and external validity. Internal validity is the basic minimum without which any experiment is uninterpretable: Did in fact the experimental treatments make a difference in this specific experimental instance? External validity asks the question of generalizability: To what populations, settings, treatment variables, and measurement variables can this effect be generalized? Both types of criteria are obviously important, even though they are frequently at odds in that features increasing one may jeopardize the other. While internal validity is the sine qua

(continued)

non, and while the question of external validity, like the question of inductive inference, is never completely answerable, the selection of designs strong in both types of validity is obviously our ideal.”

The four main considerations of external validity are as follows:

- Can the results from the typically small number of subjects (for example, persons, companies) be transferred to corresponding populations? The answers to such questions usually lie in the tools of random sampling theory and inferential statistics (see Sect. 6.4).
- Is the generalization of the results possible with regard to the object of investigation (e.g., attitude to a product → attitude to a retailer)?
- Are the results transferrable to other contexts (for example, other cultural environments, other times)?
- Does one get the same results when using other methods of examination (such as other measurements) or do the results depend on the method?

The sources of danger for the external validity of experiments are (Shadish et al. 2002):

- *Biased selection.* Selecting participants in a way that they are not representative of the population under investigation weakens the generalizability of the results.
- *Reactivity of the experiment.* The manipulations in a controlled laboratory environment may not apply to a less controllable real environment.

With regard to practical issues, external validity is indispensable, because it is about making inferences from the results of a study on the events in broader contexts (e.g., markets) for which decisions are to be made (Calder et al. 1982). This also shows that the use of experiments is by no means limited to the examination of causal relationships in theories. Particularly in practice, questions often arise such as, “What would happen if?”. The representative selection of test subjects (analogous to the typical procedure for representative surveys) and a realistic (“natural”) examination situation obviously have special significance for the external validity. However, as discussed above, these two issues often present challenges to internal validity, where homogeneity of subjects and artificial testing situations are favored to minimize the influence of confounding factors. In the literature, there are extensive discussions on how to try to increase the realism of experiments without reducing the credibility of the results, i.e., to ensure external and internal validity at the same time (Geuens and Pelsmacker 2017; Morales et al. 2017). These include, above all, the design of realistic experimental stimuli, the use of behavioral variables as dependent variables, and the composition of the sample. Because there is a *trade-off between the internal and external validity of experiments*, achieving both goals at the same time is a challenging task and almost impossible to achieve.

8.3.3 Quasi-experiments

Typical for the above-identified experimental designs are the controlled (or manipulated) use of the independent variable and the random assignment of subjects to experimental and control groups. The aim is to eliminate systematic differences between these groups that might bias the effect of the independent variables. There are situations in which these conditions do not occur. Two examples may illustrate this problem:

- To investigate whether the children of smokers are more likely to become smokers than other people: it is obvious that a random assignment to the two groups to be compared (“parents are smokers” and “parents are non-smokers”) is not only practically impossible, but also ethically highly questionable.
- To investigate whether home ownership affects budget allocation and consumer behavior over the long term (10 years or more): one will barely have 10 years to observe the consumer choice behavior of homebuyers in contrast to tenants. It would be more viable to find out from current homeowners and tenants what behavioral differences arise. That would certainly not be a random assignment, but would solve the problem of the duration of the study.

Campbell and Stanley (1963, p. 34) speak of **quasi-experiments** in situations in which essential principles of experimental investigations are applied without being able to meet all relevant requirements. There are a number of reasons for the necessity and application of quasi-experiments:

- A *randomized assignment* of subjects to the experimental groups is often *not possible*, for example, if one wants to check the effects of different viral infections.
- *Ethical reasons* often also speak against experimental manipulations, even if it were possible, such as in reviewing the effects of illegal drugs.
- The *duration of the experiment* can be too long to apply a classical experimental design, for example, in examining the long-term impact of the media on a society’s values.

Quasi-experiments thus are characterized by the fact that a randomized assignment of subjects to the experimental groups is not possible; that an independent variable cannot be manipulated and that there are no interventions that influence the dependent variable of the study.

Campbell and Stanley (1963, p. 34) on quasi-experiments:

“There are many social settings in which the research person can introduce something like experimental design into his scheduling of data collection

(continued)

procedures (e.g., the when and to whom of measurement), even though he lacks the full control over the scheduling of experimental stimuli (the when and to whom of exposure and the ability to randomize exposures) which makes a true experiment possible.”

Kerlinger and Lee (2000, p. 536) identify the reasons for carrying out quasi-experiments:

“The true experiment requires the manipulation of at least one independent variable, the random assignment of participants to groups, and the random assignment of treatments to groups. When one or more of these prerequisites is missing for one reason or another, we have a compromise design. Compromise designs are popularly known as quasi-experimental designs.”

In quasi-experiments—by the necessary absence of the random assignment of study subjects to experimental and control groups—a *confounding and distorting effect cannot be excluded*, so other ways are necessary to assure the absence of alternative explanations. Shadish et al. (2002, p. 105) emphasize the “identification and study of plausible threats to internal validity” by critically examining potential alternative influencing factors, which are typically considered as additional *control variables* in data analysis. If, for example, one wants to check whether the (non-) smoking behavior of the parents has an influence on whether the children become smokers, then it makes sense to also include control variables that describe the social environment, or the children’s personality, and provide alternative explanations. On the other hand, quasi-experiments often have advantages in terms of external validity, because the data were collected in “natural” situations.

8.4 Complex Causality

Causal hypotheses, as well as the analytical procedures for investigating causality, usually assume causal relationships that assume the necessary and sufficient conditions for an effect (for example, “the more investment, the more revenue”). **Complex causality** means distinguishing between *different forms of causality* by distinguishing between combinations of necessary and sufficient conditions. Schneider and Eggert (2014) illustrate four forms of causality, exemplifying the relationship between the two concepts of commitment and trust in a business relationship. This research assumes that trust leads to commitment in a business relationship, that is, trust is a cause, and commitment is the effect:

- One variable is a necessary but not sufficient condition for the occurrence of another variable. That is, commitment occurs when trust occurs, but does not need to, so that trust can occur without there being any commitment.

- A variable is a sufficient but not a necessary condition for a second variable. That is, commitment occurs when trust occurs, but commitment can also occur without trust.
- A variable can be part of a combination of sufficient conditions without itself being sufficient or necessary. Trust might explain commitment sufficiently well, but only in combination with other factors, such as the benefit of a relationship. Trust would then be a so-called INUS condition (see Sect. 8.1).
- One variable is a sufficient and necessary condition for the occurrence of a second variable. That is, trust always leads to commitment and commitment without trust does not occur.

The typical technique used to analyze complex causalities is **Qualitative Comparative Analysis (QCA)**. QCA is a method of causal analysis of configurational data in the social sciences. Configuration data means that all variables, no matter what measurement levels, are converted to qualitative data, for example, different levels of trust, which are typically measured as an interval-scaled variable, convert to “trust exists/trust does not exist”. Furthermore, there is a difference between an “outcome”, which in principle is the effect (here: commitment), as well as the “conditions”, these are the causes and possible moderators (here: trust, benefit of a relationship, etc.). For each observation (e.g., for each business partner), a value between 0 and 1 is entered into a truth table for the conditions and the outcome, which indicates to what extent the observation tends towards one or the other characteristic of the configurational variables (e.g., the probability of the occurrence of trust or commitment). Subsequently, algorithms are applied, with the search objective to identify minimally necessary and sufficient conditions for the presence of the outcome: if, for example, in all observations in which commitment (the outcome) is found, there is always trust, then trust is a necessary condition for commitment. For the details of this analysis, please refer to the relevant literature (e.g., Ragin 2008; Schulze-Bentrop 2013). The result of the analysis indicates those conditions that are necessary and those that sufficiently explain the outcome. This can be a single condition, but it can also be combinations of conditions.

The advantage of QCA over other, non-experimental methods of causal analysis is the identification of the causes of an effect. However, if one wants to examine how much one particular variable (cause) contributes to the explanation of another variable (effect), then conventional regression-based analysis techniques are more appropriate.

References

- Arabatzis, T. (2008). Experiment. In S. Psillos & M. Curd (Eds.), *The Routledge companion to philosophy of science* (pp. 159–170). London: Routledge.
- Aronson, E., Wilson, T., & Brewer, M. (1998). Experimentation in social psychology. In D. Gilbert, S. Fiske, & G. Linzay (Eds.), *The handbook of social psychology* (4th ed., pp. 99–142). New York: McGraw-Hill.
- Bagozzi, R. P. (1980). *Causal models in marketing*. New York: Wiley.

- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.
- Calder, B., Phillips, L., & Tybout, A. (1982). The concept of external validity. *Journal of Consumer Research*, *9*(3), 240–244.
- Campbell, D. T., & Stanley, T. D. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Cartwright, N. (2014). Causal inference. In N. Cartwright & E. Montuschi (Eds.), *Philosophy of social science* (pp. 208–326). Oxford: Oxford University Press.
- Chalmers, A. (2013). *What is this thing called science?* (4th ed.). Indianapolis, IN: Hackett.
- De Vaus, D. (2001). *Research design in social research*. London: Sage.
- Geuens, M., & Pelsmacker, P. D. (2017). Planning and conducting experimental advertising research and questionnaire design. *Journal of Advertising*, *46*(1), 83–100.
- Godfrey-Smith, P. (2003). *Theory and reality—An introduction to the philosophy of science*. Chicago: University of Chicago Press.
- Hitchcock, C. (2008). Probabilistic causation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <http://plato.stanford.edu>
- Humphreys, P. (2000). Causation. In W. Newton-Smith (Ed.), *A companion to the philosophy of science* (pp. 31–40). Malden, MA: Wiley.
- Jaccard, J., & Becker, M. (2002). *Statistics for the behavioral sciences* (4th ed.). Belmont, CA: Wadsworth.
- Jaccard, J., & Jacoby, J. (2010). *Theory construction and model-building skills—A practical guide for social scientists*. London: Guilford.
- Kerlinger, F., & Lee, H. (2000). *Foundations of behavioral research* (4th ed.). Melbourne: Wadsworth.
- Koschate-Fischer, N., & Schandelmeier, S. (2014). A guideline for designing experimental studies in marketing research and a critical discussion of selected problem areas. *Journal of Business Economics*, *84*(6), 793–826.
- MacKenzie, S. B., Lutz, R. J., & Belch, G. E. (1986). The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations. *Journal of Marketing Research*, *23*(2), 130–143.
- Morales, A. C., Amir, O., & Lee, L. (2017). Keeping it real in experimental research—Understanding when, where, and how to enhance realism and measure consumer behavior. *Journal of Consumer Research*, *44*(2), 465–476.
- Mumford, S., & Anjum, R. (2013). *Causation—A very short introduction*. Oxford: Oxford University Press.
- Peter, J. (1991). Philosophical tensions in consumer inquiry. In T. Robertson & H. Kassirjian (Eds.), *Handbook of consumer behavior* (pp. 533–547). Englewood Cliffs, NJ: Prentice-Hall.
- Psillos, S. (2002). *Causation & explanation*. Durham: Acumen.
- Psillos, S. (2004). Causality. In M. Horowitz (Ed.), *New dictionary of the history of ideas* (pp. 272–280). New York: Charles Scribners.
- Ragin, C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago: University of Chicago Press.
- Schneider, M. R., & Eggert, A. (2014). Embracing complex causality with the QCA method: An invitation. *Journal of Business Market Management*, *7*(1), 312–328.
- Schulze-Bentrop, C. (2013). *Qualitative comparative analysis (QCA) and configurational thinking in management studies*. Frankfurt am Main: Peter Lang.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Further Reading

- Geuens, M., & Pelsmacker, P. D. (2017). Planning and conducting experimental advertising research and questionnaire design. *Journal of Advertising*, *46*(1), 83–100.
- Koschate-Fischer, N., & Schandelmeier, S. (2014). A guideline for designing experimental studies in marketing research and a critical discussion of selected problem areas. *Journal of Business Economics*, *84*(6), 793–826.
- Psillos, S. (2002). *Causation & explanation*. Durham: Acumen.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.