Judgment and Decision Biases

Part I & II

Author: Siegfried Macho

Date of last modification: *Tuesday, October 29, 2019*

Table of Contents

1. Introduction: Basic Considerations

The present introductory chapter presents basic concepts and principles that are relevant for the subsequent treatment of the topic.

1.1 Judgmental Errors and Human Rationality

In the following we regard judgmental errors as one form of irrational behavior. This raises the following questions:

Questions:

1. *How could we assess whether a judgmental error is present or not?*
2. *What do we mean by rational and irrational behavior?*

In answering the first question consider the following principle:

Principle 1-1: *Identification of judgmental errors:*

A (human) judgment may be classified as erroneous if the following conditions are satisfied:

1. *There exists a normative principle (rule).*
2. *This rule can be applied to the actual situation.*
3. *The reasoning person violates this normative principle.*

Previously to discussing a concrete example let us clarify what is meant by the term *rational behavior*.

Principle 1-2: *Rational behavior:*

Human (and animal) behavior may be considered as rational if it maximizes the expected subjective utility.

or, stated differently:

Human behavior may be considered as rational if it enables the attainment of personal goals (whatever this goal may be) in an optimal way.

Let's consider a simple illustrative example.

Ex. 1-1: Transitivity of preferences:

An important principle of rationality states that personal preferences should be transitive. Specifically, the principle states:

If a person prefers Alternative A to Alternative B, and if, at the same time, she prefers B to Alternative C then she should prefer A to C.

In symbols:

$$A \succ B \wedge B \succ C \rightarrow A \succ C$$

where:

$A \succ B$ means that Alternative *A* is preferred to *B*;

The symbol $\wedge$ represents the logical AND: the statement $X \wedge Y$ is true if and only if *X* and *Y* are both true.

The symbol $\rightarrow$ represents the logical implication: the statement $X \rightarrow Y$ is true if and only if *X* is wrong or *Y* is true.

It is quite simple to demonstrate that a person violating the principle of transitivity can be made a »money pump«, that is, one can construct a series of transactions such that the person loses permanently money.

In order to demonstrate this consider the following situation:

Consider three different text books in statistics: »The Bortz« (*B*), »the Freedman« (*F*), and »the Hayes« (*H*).

Person *P* holds the following intransitive preference ordering:

1. $B \succ F$
2. $F \succ H$
3. $H \succ B$

Now, assume that I am in possession of one of three books, say *B*. Our target person, call her *P*, owns the other two books: *F* und *H*.

Since *P* prefers *B* to *F*, she will be willing to pay a small amount of money, say Sfr. 1.- (the exact amount of money is not important), to exchange *B* for *F*.

Similarly, she will be willing to pay a certain amount of money (again, say, 1.-), to exchange *H* for *B* (since she prefers *H* to *B*).

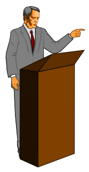
Finally, she should accept to pay a small amount of money (again, say Sfr 1.-) to exchange *H* for *B* (due to her preference of *H* over *B*).

Thus, we are back at the starting configuration (with *P* having lost Sfr. 3.-), and the next round of game can begin (Tab. 1-1 illustrates the sequence of transactions and the preferences used in each round).

Round	Me	Person P	Preference
1.	B	F, H	$B \succ F$
2.	F	B, H	$F \succ H$
3.	H	B, F	$H \succ B$
4.	B	F, H	

Tab. 1-1: Sequence of transactions with intransitive preferences:

This simple example illustrates that Person P harms herself by violating the principle of transitivity of preferences. She thus does not maximize her expected personal utility. By consequence, she may be regarded as acting irrational.

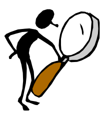


Comment 1-1: The basic axioms of probability theory and the »Dutchbook« argument

The axioms of probability theory (cf. Appendix) can be justified by similar arguments. Specifically, it can be shown that in case of violating an axiom a system of bets, called a *»Dutchbook«*, can be constructed that lead to certain loss.

The system consists of bets only that, when considered in isolation, are perfectly fair bets (for a simple exposition of the argument, cf. Hacking, 2001)

The formulation of Principle 1-2 comprises an »intricacy« that should be noticed: It does not state that the personal utility has to be maximized. Rather, it states that the *expected* personal utility should be maximized.
To understand this subtle difference we first define the concept of expected subjective utility.



Concept 1-1: *Expected subjective utility:*

Given:

- ❑ Possible consequences $C_1, C_2, \ldots, C_m$, resulting from the choice of an alternative.
- ❑ Subjective utilities $u_1, u_2, \ldots, u_m$ associated with the single consequences where the utility u_i resulting from a specific consequence C_i may be positive or negative.
- ❑ Probabilities $p_1, p_2, \ldots, p_m$ that the single consequences occur: $P(C_i) = p_i$.

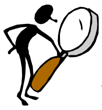
The *subjective expected utility* $SEU(X)$ of an Alternative X corresponds to the sum of the single subjective utilities multiplied by the corresponding probabilities:

$$SEU(X) = u_1 \cdot p_1 + u_2 \cdot p_2 + \cdots + u_m \cdot p_m$$

The Subjective Expected Utility (SEU) is used as the objective to be maximized since the actual utility resulting from the decision depends on chance. By consequence, it is possible that a person makes a bad choice with respect the expected utility. Yet, due to good luck, her choice results in a high subjective utility. On the other hand, a person may decide optimally with respect to expected utility. However, due to misfortune her decision results in a low subjective utility.
In order to exclude this dependence of an actual outcome on pure chance (that cannot be controlled by person) the relevant quantity is the subjective expected utility.

1.1.1 Epistemic vs. Instrumental Rationality

On the basis of the previous presentation the presence of judgmental (and decision) biases may be assessed using two different types of criteria: (i) The judgment is in contradiction to normative principles, and (ii) the judgment or decision prevents the optimization of the expected personal utility. In the current literature different terms are used to refer to the two concepts (Stanovich, 2010).



Concept 1-2: *Epistemic vs. instrumental rationality:*

Epistemic (theoretical) rationality is realized if the judgment and decision, respectively, is in accordance with principles that are accepted as being normatively correct (and these principles apply to the current circumstances).
Instrumental (practical) rationality is accomplished if the judgment and decision, respectively, optimizes the subjective expected utility.

This raises the following question:

Question:
Do epistemic and instrumental rationality necessarily result in the same judgments and decisions?

The answer to this question is not straightforward. In fact, there is a debate about rationality in psychology comprising different points of view (cf. Stainton, 2006; Stanovich, 2010).
However, the following example indicates that both types of rationality may well lead to different outcomes.

Ex. 1-2: Well-being and the belief in a life after death

The belief in a life after death contradicts an approved principle of rationality, called *Occam's Razor* [named after the English monk and scholastic philosopher William of Ockham (presumably 1288-1347)].

According to this principle entities should not be proliferated without necessity, that is, the existence of an entity should not be taken for granted if there is no definite reason or evidence (»Entia non sunt multiplicanda praeter (sine) necessitatem«).

According to this principle the belief in a life after death is not justified since there is no evidence of a life after death. Consequently, this belief violates the principles of epistemic rationality.

On the other hand, the belief in a life after death might maximize expected personal utility since there are a great deal of studies from the domain of *Positive Psychology* demonstrating that humans who belief in a life after death enjoy themselves of a higher psychological well-being. In addition, they exhibit a higher ability to master blows of fate.

By consequence, the belief in a life after death may be well justified with respect to instrumental rationality (Assuming that the gain in well-being trades off the costs that are associated with a belief in an afterlife, like the consciousness of sins and the coercion to execute certain rituals).

Comment 1-2: On the importance of Occams razor

Occam's razor has exerted a great impact on human's history: It resulted in the demystification of the world relieving it from different kinds of super natural entities, like ghosts and spirits.

Also in science Occam's razor plays an important role since it implies that simple theories have to be preferred to more complex ones that make more assumptions.

For example, the fact that the null hypothesis H_0 is rejected only if the probability of the observed data given H_0 is low ($p < .05$) is a direct consequence of Occam's razor since the alternative hypothesis is, in general, the more complex one.

Comment 1-3: Final comment on different types or rationality

The two types of rationality (epistemic and instrumental) comprise the most important types of rationality. In the literature one may find further types of rationality. For example, Kahneman und Frederick (2002) use the term *coherence rationality*, meaning that a person's beliefs have to be coherent (e.g. personal probabilities have to conform to Bayes theorem). However, this type of rationality is but a subtype of epistemic rationality.

1.1.2 Possible criticisms concerning the research of judgmental and decisional biases as well as the concept of rationality

In recent years, the traditional research on judgmental and decision biases has been subjected to a rigorous criticism that took three different directions:

1. *Criticism of normative principles:*
   This type of criticism concerns the adequacy of the assumed normative principles that are used as a standard for evaluating the adequacy of judgments and decisions. For example, Cohen (1981) argues that normative principles are conventions and thus nothing more than subjective assumptions.
2. *Criticism of the applicability of normative principles:*
   According to this criticism normative principles used as a standard for assessing participants' performance in various investigations are not applicable or they were applied in the wrong way (see, for example, Birnbaum, 1983; Gigerenzer, 1996; Gigerenzer & Murray, 1987).
3. *Problems of ecological representativeness:*
   According to this type of criticism the biases found in experiments are artificial. They are due to fact that the tasks presented to participants do not adequately reflect reality (see, for example, Gigerenzer & Hoffrage, 1995; Juslin, 2001).

These criticisms will be discussed in due course. Specifically, in Chapter 0 the first two types of criticisms will be addressed. However, as an introduction into the problem area, the following example illustrates the issue of the applicability of normative principles.

Ex. 1-3: Expected utility of singular events:

Assume that someone offers you the following bet:

A dice will be thrown. If the number of points is 5 or 6 you get Sfr. 6000.- otherwise you pay Sfr. 600.-

Most people will reject this bet despite the fact that its expected utility is positive:

$$E_U = \frac{1}{3}\cdot 6000 - \frac{2}{3}\cdot 600 = 1600$$

The reason for the rejection is ostensibly due to the fact that for most people a loss of Sfr. 600.- is rather painful. In addition, the probability of a loss: $P(loss) = 2/3$, is higher than that of a gain: $P(gain) = 1/3$.

One might argue that the concept of *probability* does not apply to singular events (at least according to the frequentist theory of probability). By consequence the computed expected utility makes little sense.

This is in accordance with the fact that most people would accept the following modified version of the bet:

A dice will be thrown. If the number of points is 5 or 6 you get Sfr. 600.- otherwise you pay Sfr. 60.-

The game will be repeated 10 times and the gains and losses from the single trials will be added.

This version of the bet has exactly the same expected utility as the original one. It is however, much more attractive for most people.

Note that with the modified version of the bet one can win Sfr. 60.- if in at least one trial the number 5 or 6 appear. The probability of this event is quite high:

$$P(At\ least\ one\ gain) = 1 - \left(\frac{2}{3}\right)^{10} = 0.983$$

There is a second aspect that is related to the first one: The variance of the second bet is 10 times lower: 968000 vs. 9680000 (for the original version).

In conclusion one might argue as follows:

The seemingly irrational behavior of most people to reject the original form of the bet turns out, on closer inspection, as an erroneous application of the concept of probability.

Or, alternatively:

In their choices, people do not only take the expected utility into account but consider also the variance of the utility. It does not make sense to indicate persons that also consider the variance of the utility as irrational.

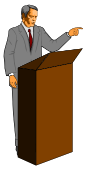
Interestingly, the idea that the variance should be taken into account has found little attention in the literature. One exception is Rode, Cosmides, Hell, & Tooby (1999).

1.2 Psychological Mechanisms: »Hot« und »cold cognition«

Social psychology has afforded a distinction between »hot« und »cold cognition«. The label »hot cognition« is due to Abelson (1963).

Concept 1-3: *Hot vs. cold cognition:*

The term *hot cognition* refers to cognitive processes (e.g. judgment and decision processes) that are strongly influenced by such factors as motivation, emotion, mood or states of arousal (e.g. stress).
In case of *cold cognition* these factors are virtually absent and thus exert no (or little) influence on the actual cognitive processes.



Comment 1-4: Cold blood

The differentiation between hot and cold cognition reflects a similar distinction that is of great importance in jurisdiction, namely, whether one has acted in a state of high arousal or in cold blood.

Clearly, the distinction between hot and cold cognition should be conceived of as gradual one since judgments and decisions have some sort of motivational basis. For example, most people are, in general, interested to provide a good judgment or decision.
In talking about the influence of motivational factors on judgments and decisions one usually means that the motivational factors exert some sort of force to bias the judgment *in a certain direction* in order to avoid cognitive dissonance (Kunda, 1990).
Typical examples of this type of biased judgments are:

- ❑ Misattributions in order to avoid a negative self-image: »It was not me who was aggressive but my behavior was provoked by the other guy«.
- ❑ Historical lies: Holocaust denial;
- ❑ Conspiracy theories.
- ❑ Denial of the biological evolution;
- ❑ Ignoring or misinterpreting empirical evidence;
- ❑ Ignoring approved scientific knowledge.

However, in some cases both hot and cold cognition may be involved as in some instances of *myside bias*. The latter consists of bias in generating of evidence, testing of hypotheses and evaluation of policies directed toward the own opinion.

Ex. 1-4: Myside bias:

Consider the following two scenarios presented to two different groups of Swiss participants:

Myside (Swiss version):

The Swiss army plans to buy 300'000 new army knives. The evaluation of various offers from different providers arrives at the conclusion that a Chinese company provides the most favorable offer concerning the balance between price and quality.
Following to a great protest within Swiss public the Swiss army decides to buy the knives of a Swiss provider (Victorinox).
How would you evaluate the decision of the Swiss army to buy the knives of the Swiss provider?
6-Point Scale: 6 = definitively right, 1 = definitively wrong.

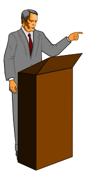
Otherside (Chinese version):

The Chinese army plans to buy 300'000 new army knives. The evaluation of various offers from different providers arrives at the conclusion that the Swiss company Victorinox provides the most favorable offer concerning the balance between price and quality.
Following to a great protest within the Chinese public the Chinese army decides to buy the knives of a Chinese provider.
How would you evaluate the decision of the Chinese army to buy the knives of the Chinese provider?

Myside bias is exhibited if participants receiving the Swiss version indicate higher agreement compared to participants getting the Chinese version.

An observed myside bias may be in part due to patriotic sentiments [hot cognition]. On the other hand it demonstrates a certain degree of an inability to take the view point of the other side (egocentristic bias, lack of open-mindedness) [cold cognition].

In the following we shall be mainly (but not completely) concerned with judgmental and decisional errors and biases, respectively that are *not* due to motivational factors. Rather, the biases discussed are predominately due to basic cognitive constraints, i.e. a fundamental inability to correctly represent and process the relevant information.
A good overview of the issue of biased judgment due to motivational causes is given by Kunda (1990). Janis und Mann (1979) present investigations concerning decisions under stress or influenced by other motivational factors.



Comment 1-5: Bias and motivation

In many cases, the presence of motivational factors results in a small bias only. This is due to the fact that people seek to behave rational thus trying to avoid judgmental errors (Kunda, 1990).

In the following chapters we investigate judgmental and decision errors in different domains. The principal empirical results will be presented first. This will be followed by an explication of the underlying psychological mechanisms. Finally, concepts and methods will be presented that enable one to adequately analyze the situation and to find correct answers.

2. Contingency and Causal Judgments

Humans' history is paved with examples of flawed judgments about associations between events (= contingency judgments) and of faulty judgments about cause-effect relationships (= causal judgments). On the one hand there are cases where spurious (non-existing) [causal] relationships have been postulated, and on the other hand important [causal] relations between events have not been detected.
Here are a number of noteworthy examples of *spurious contingency judgments*:

- Astrology: Associations between constellation of planets at the time of birth and personal characteristics.
- Pseudo-scientific diagnostic procedures, like graphology, lie-detectors, projective tests (e.g. Rorschach test, draw a man test, thematic apperception test etc.).
- Diagnostic procedures found sometimes in magazines, like tests that enable the prediction of personality characteristics on the basis of sleep position, dietary habits, furnishing, etc.
- Ernst Kretschmer's (1988-1964) typology postulating an association between body composition and character.

Concerning examples of spurious causal judgment, many examples can be found in Psychology and medicine:

- A negative effect of full moon on human's behavior.
- A negative effect of traumatic events on memory (repression).
- A positive effect of post traumatic talks on the processing of the traumatic event.
- Positive effects of specific treatments, like homeopathy or acupuncture, on health.
- Positive effects of treatments on health that in fact have turned out to exert a negative effect on health: arteriotomy (cf. Ernst & Singh, 2009).

An example of the failure to detect an important causal relationship is the underestimation of the significance of hygiene for preventing death in hospitals. At first, the pioneering work of Florence Nightingale (1820-1910) and Ignaz Semmelweis (1818-1865) was met with great skepticism.
In order to understand the psychological mechanisms underlying defective contingency and causal judgments let us first discuss a classical empirical study on this topic.

2.1 A Classical Study of Chapman and Chapman on Illusory Correlation

In 1969 Chapman and Chapman conducted a classical empirical study on a phenomenon called *illusory correlation*. The study used the Rorschach test.

The Rorschach test comprises 10 cards with inkblots (some of the cards are only in black and white and some are colored). The client receives a card with the instruction to »identify« shapes in figures in the inkblots.

It turned out that, compared to heterosexual men, homosexuals »perceive« more frequently monsters on Card 4. Consequently, the identification of monsters on Card 4 can be considered as a valid indicator of homosexuality. There are still other valid indicators of homosexuality.

Additionally to these valid signs there are a number of *plausible* indicators (= indicators with high face value), for example, identified persons with female cloths, persons with unclear sex, identified masculine genitals or the perception of anal shapes.

The latter indicators possess a high face value since common sense suggests that homosexual men might exhibit an increased tendency to project these kinds of shapes into the inkblots. However, there exists no statistical association between plausible indicators and homosexuality.

The investigation of Chapman and Chapman (1969) comprised four stages:

Stage 1:

32 practicing psychiatrists who had indicated to have analyzed Rorschach protocols of a number of homosexual men received a list of indicators (i.e., the answers of hypothetical clients). These protocols contained valid but implausible as well as invalid but plausible indicators of homosexuality.

The five most frequent answers that were classified as the most relevant by the clinicians concerned invalid but plausible indicators. Only two clinicians mentioned the empirical valid indicators.

Stage 2:

Naïve participants were asked to rate the degree of association between homosexuality and the various indicators. Unsurprisingly, the relationship between invalid indicators with high face value was rates as »relatively strong« whereas the association between the valid but implausible indicators and homosexuality was judged being as »quite weak«.

Stage 3:

Naïve participants were trained to learn the association between homosexuality and indicators, i.e. they should learn which answers on the Rorschach test are valid indicators of homosexuality and which are not.

Symptoms were either of the homosexual type (e.g., »He feels attracted by other men«) or neutral with respect to homosexuality (e.g. »He often feels lonely«).

Participants were presented 30 Rorschach cards with the inkblots on one side and the answers as well as symptoms of the purported clinical subject on the other side. The answers of the clinical subjects consisted of valid, invalid but implausible, and invalid but neutral answers.

The series of 30 cards was constructed in such a way that there was no association between the answers (of any type) and homosexual symptoms.

In each trial participants received a single card and had 60 seconds to study the information. Following the presentation of the 30 cards participants had to indicate which answers were associated with homosexual symptoms.

Despite the fact that there was no objective relationship between plausible answers and homosexual symptoms participants perceived an association, similarly to the »experienced« clinicians.

Stage 4:

The strength of the association between valid answers and homosexual symptoms was varied: The relationship was 50%, 67%, 83%, or 100%.

The real strength of the association had no influence on participants' judgments: They were unable to detect the relationship between valid answers and homosexual symptoms.

Only after removal of the plausible answers from the cards participants were able to detect the correct relationships.

The main results of the study of Chapman and Chapman (1969) may be summarized as follows:

1. Due to their preconceptions »experienced« clinicians as well as naïve people detected non-existing associations (= *illusory correlation*).
2. The same preconceptions prevent people to learn valid but implausible association.

2.2 Subjective Theories and Erroneous Contingency and Causal Judgments

The investigations of Chapman und Chapman (1969) indicate an important mechanism underlying erroneous contingency and casual judgments.

Cognitive Mechanism 2-1: The significance of subjective theories on the perception of relationships between events

Faulty subjective (plausible) theories about (causal) relations can exhibit an negative effect on the contingency of causal judgment in the following way:

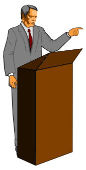
(i) They lead to the »perception« of non-existing relations.

(ii) They prevent the recognition of existing associations.

As will be illustrated in the following sections, erroneous subjective theories are not only responsible for defective contingency and causal judgment but also for another important bias, namely the biased processing of empirical evidence.
The given explanation raises the following question:

***Question**:*
How do people acquire erroneous subjective theories?

Most theories about contingency and causal relations are culturally conveyed. Thus, most erroneous subjective theories are culturally acquired. Specifically, so called »experts« constitute an important source of faulty subjective theories. However, these authorities are often experts in quite a different field.
A nice example is Linus Pauling (1901-1994) who received two noble prices (chemistry and piece). On the basis of a defective study he claimed that the increased consumption of Vitamin C (ascorbic acid) protects from cancer.
Detailed investigations of the Mayo clinic did not confirm this claim (Ironically, Linus Pauling himself died from cancer).



Comment 2-1: Psychology as a source of erroneous theories:

Lilienfeld, Lynn, Rusco und Beyerstein (2010) discuss many examples of erroneous psychological theories or hypotheses, like:

- Traumatic events are repressed;
- Learning of new material during sleep is possible;
- Anger should be acted out.

Comment 2-2: Neuropsychology as a source of erroneous theories

Macdonald, Germine, Anderson, Christodoulou, & McGrath, (2017) present a list of beliefs in neuromyths, e.g.:

- We use only 10% of our brain;
- Some of us are »left-brained« and some are »right-brained« and this explains differences in how we learn;
- When we sleep, the brain shuts down.

The so called »common sense« constitutes a second source of defective subjective theories. It is based on, apparently obvious or self-evident facts (Einstein once uttered: »According to common sense the earth is a disk«). According to common sense the similarity between cause and effects seems to be an important criterion in causal judgment. Gilovich und Savitsky (2002) present a number of examples from different domains demonstrating the significance of similarity for causal and contingency judgments.

Ex. 2-1: Significance of similarity for causal und contingency judgments

Medicine:

- Homeopathy: Use the same drug that caused the disease.
- In case of problems of virility use the pulverized horn of rhinoceros (it is well-known that the horn of the rhino is tough).
- In case of asthma use the lungs of foxes (foxes are well-known for their endurance).
- Microorganisms are much too small to do great harm (since only big things can do great harm).
- In case of damage of specific organs one should eat the same organ of wild beasts raw and in natural simplicity. For example, eat parts of the brain in case of mental diseases, or raw stomachs in case of stomach ulcer.

Negative effects of foods:

- Fatty potato chips result in oily skin.
- Highly seasoned foods lead to hart burning.

Astrology:

- Twins are double-faced (ambiguous personalities).
- Capricorns are adamantine and stubborn.
- Lions are pride and resolute leaders.
- Cancers have got a hard shell and a soft core.
- People born in the star sign of Libra have a balanced and harmonic personality.

Graphology:

- Zonal theory: The upper part of the handwriting reveals intellectual characteristics, the middle part practical and the lower part aspect of the instinct.
- The distance of the signature from the written text reveals how strongly the person distances herself from the written text.
- If the handwriting remains near the left border then the person stay stuck in the past. If the right border is included too the person is focused on the future.

An important characteristic of folk theories on which common sense is based consists in their vagueness. By consequence folk theories are difficult to refute. The fact that people accept vague explanations is nicely demonstrated by the so called Barnum effect.

Concept 2-1: *Barnum Effect (Forer Effckt):*

Der *Barnum Effect* – derived from the American entertainer Phineas Taylor Barnum (1810-1891) – describes the tendency of people to assess superficial and vague descriptions of the own personality as quite accurate.

In 1948 the psychologist Bertram R. Forer presented to people a description of their personality. The latter had to rate, on a scale from 0 to 5 (very accurate), how well the description fitted the own personality.

Here is the description:

You have a great need for other people to like and admire you. You have a tendency to be critical of yourself. You have a great deal of unused capacity which you have not turned to your advantage. While you have some personality weaknesses, you are generally able to compensate for them. Disciplined and self-controlled outside, you tend to be worrisome and insecure inside. At times you have serious doubts as to whether you have made the right decision or done the right thing. You prefer a certain amount of change and variety and become dissatisfied when hemmed in by restrictions and limitations. You pride yourself as an independent thinker and do not accept others' statements without satisfactory proof. You have found it unwise to be too frank in revealing yourself to others. At times you are extroverted, affable, sociable, while at other times you are introverted, wary, reserved. Some of your aspirations tend to be pretty unrealistic. Security is one of your major goals in life.

Each of the participants received the same description. The mean rating was: 4.26.

Employed strategies:

The different statements in the foregoing passage may be classified into various categories where one statement may belong to more than one category:

1. Truisms that apply more or less to any human being. Consequently they are of little diagnostic value:

   You have a great need for other people to like and admire you.

   Since (nearly) any human being has a desire to be admired and liked the passage is quite meaningless.

2. Flattery:

   You have a great deal of unused capacity which you have not turned to your advantage.

   Statements like this one flatter our vanity and are thus well accepted.

3. Statements that are partly true and enable different interpretations:

   At times you are extroverted, affable, sociable, while at other times you are introverted, wary, reserved.

4. Statements with restrictive remarks:

   While you have some personality weaknesses, you are generally able to compensate for them.

   The generality of the statement is limited by additional remarks.

5. The feeling to be seen through by an expert.

The problem of vagueness and ambiguity not only concerns folk theories but also »scientific theories« in (neuro-) psychology and other branches of the humanities.

Ex. 2-2: False pretenses of precision:

Aron, Robbins, and Russel (2004, 2014) argue that the right inferior prefrontal cortex (rIFC) is crucial for cognitive control. However, in reading the article one gets the impression that in most tasks the are considering a great part of the right and left prefrontal cortex is involved:

The subregion of the IPFC most commonly implicated is the pars opercularis [...]; however, there is often activation of the right (and left) insula, pars orbitalis, and triangularis and the inferior frontal junction [...] (Aron et al, 2014, p.177).

Finally, the authors conclude that the rIFC as well as a »*wider prefrontal-basal network«* (Aron et al, 2014, p.183) function as a sort of cognitive brake.
In the abstract and conclusion of the article, the authors pretend a precision with respect to the localization of cognitive functions that seemingly does not exist.

The vagueness of subjective theories may be one important cause for the arbitrariness of folk explanations of human behavior.

Ex. 2-3: Arbitrariness of everyday explanations:
(Ross, Lepper, Strack, & Steinmetz 1977):

Participants were asked to place themselves in the position of a clinical psychologist attempting to understand and predict patient's performance on the basis of specific background information.
The background information consisted of the authentic clinical case histories. Participants were asked to try to explain critical life event of the patient (e.g. a suicide, joining a Peace Corp, a hit-and-run accident, a candidacy for a political office) on the basis of the presented information.

Result:

Participants had absolutely no problems to explain the critical events. However, as noted by Nisbett and Ross (1980), the explanations exhibited a horrendous amount of arbitrariness.

> *»An analysis of the actual explanations written by subjects was instructive. The facility with which subjects could pass from almost any real event in the past history of the client to almost any hypothetical event was positively alarming. For example, one patient's youthful decision to join the Navy was cited as very significant both by subjects asked to explain a subsequent candidature for political office and by subjects asked to explain a subsequent suicide!*
> *In the former case, however, service in the Navy was seen as symptomatic of the "gregariousness" and the "desire to serve" that might characterize a potential politician, while in the latter case it was seen as a symptom of the patient's predisposition to "punish others by running away," thus foreshadowing a suicide.*
>
> *[**Any resemblance between the behavior of subjects in this experiment and that of any clinician living or dead, is purely coincidental.**] (Nisbett & Ross, 1980,p. 185)*

Erroneous subjective theories constitute an important cause of faulty causal and contingency judgments. There exist, however, further cognitive mechanisms that are responsible for a biased perception of associations between events.

2.3 Discovering Patterns in Random Sequences

> *[…] we are all too good at picking out non-existent patterns the happen to suit our purposes.*
> (Efron & Tibshirani, 1993, p. 1)

Humans (as well as higher animals) are experts in the processing of complex patterns. Up to now there does not exist any artificial device that achieves only remotely humans' performance in the processing of faces or complex language structures. However, the extraordinary ability to recognize complex patterns has a negative side: Humans sometimes perceive patterns even if no regularities exist.

Wagenaar (1970) presented his participants sequences consisting of white and black spots. The probability of a switch of colors between successive spots was varied in steps in .10 from .20 to .80: For one sequence of spots the probability of switching was .20, for another sequence it was .30 etc. Clearly, the switching probability of a pure random sequence with an (expected) equal number of white and black spots is .50.

Participant had to indicate that sequence which they found to be most typical of a pure random sequence with equal number of alternatives. On average sequences with a switching probability of .60 were assessed as being most typical of a pure random sequence. Thus, according to participants the rate of switching has to be higher than .50 (the correct switching probability).

Cognitive Mechanism 2-2: Representativeness and randomness

For lay people randomness is associated with a lack of orderliness and regularity, respectively.
By consequence a random sequence should not contain longer partial sequences exhibiting orderliness (like longer sequences of spots of the same color of sequences of alternating colors). Thus, each subsequence should be *representative* for the whole random sequence, i.e. they should exhibit a lack of orderliness and predictability.
Therefore, people possess an idea (a sparse subjective theory) about random processes and their outcomes. If an observed sequence does not correspond to (is not similar to) the prototype it will not be regarded as random.

According to Cognitive Mechanism 2-2 sequences of binomial events of probability .50 are not regarded as purely random since they are regarded as being too systematic. Specifically, for longer sequences the presence of subsequences comprising identical events has a high probability. However, this is in opposition to the idea that each subsequence of a random sequence should reflect the principle feature of random sequence: a lack of regularity.
The belief that random processes are represented by each subsequence of an outcome sequence may be regarded as one facet of a class of phenomena that result from a false belief that Tversky and Kahneman (1971) called the *belief in the law of small numbers.*

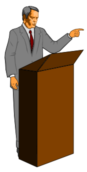
Cognitive Mechanism 2-3: Belief in the law of small numbers

The belief in the law of small numbers consists in the overestimation of the representativeness of small samples from a population, specifically:

- ❑ People overestimate the similarity between different small samples as well as the similarity between small samples and the population.
- ❑ A sample is assessed as more representative for the population if the later is smaller, that is, the relative size of the population and the sample is regarded as being significant.

  Comment: Remember that a sample consists of a series of independent random draws from the population. Consequently, the size of the population is completely insignificant with respect to the representativeness of a sample.

- Overestimation of the representativeness of subsamples of the sample: People assume that each part of a sample has to reflect the relations within the population.
- People conceive random processes as self-correcting. This inclination is called the *gambler's fallacy*.
  Comment: Random processes are not self-correcting but thinning out, i.e., within long sequences subsequences revealing a specific pattern (e.g. subsequences with the same outcome) are of low importance.



Comment 2-2:

The label *law of small number* is an allusion to the well-known *law of large numbers*. According to the latter, with increasing sample size N the sample becomes an increasingly better representation of the underlying population.
For example, with increasing N, sample statistics, like the sample mean $\bar{x}$, provide an increasingly better approximation of the corresponding population parameter.
With respect to the sample mean $\bar{x}$ this is exhibited by the fact that the standard error of the mean:

$$\sigma_{\bar{x}}^2 = \sigma^2 / N$$

decreases with increasing sample size (as well as the fact that $\bar{x}$ is an unbiased estimate of the corresponding population parameter μ, i.e. $\mathrm{E}(\bar{x}) = \mu$, where $\mathrm{E}(\bar{x})$ denotes the expectation of $\bar{x}$).

The belief in the law of small numbers reveals that lay people (as well as some experts) have developed a wrong conception of the nature of random processes. This is a further source of »detecting« non-existing relationships.
In many cases the interpretation of sportive events is affected by the belief in the law of small numbers. This is nicely illustrated by the following example (The regression to the mean and its neglect constitutes another important source of the faulty interpretation of sportive events (cf. Section 2.4)).

2.3.1 The »Hot Hand« in Basketball

Most professionals in Basketball (player, coaches) as well as many lay persons nurture the idea of the so called »hot hand« in basketball. The »hot hand« consists in the assumption that a basketball player who has scored becomes »hot«, i.e. the probability of achieving further hits increases. Gilovich, Vallone, & Tversky (1985) conducted a study to assess the validity of this assumption.

Ex. 2-4: The »hot hand in basketball«
(Gilovich, Vallone, & Tversky, 1985):

1. The authors first demonstrated that the assumption of a hot hand is quite common: Of 100 students from the Stanford and Cornell University that were interested in basketball 91 percent indicated their belief in a player's increased chance of scoring if his previous 2 or 3 shots were successful.
2. The authors evaluated the score statistics of the 9 most important players of 48 home games of the Philadelphia 76 within the season 1980-81. This revealed the following result:

   (i) Die probability of a hit following 1, 2, or 3, previous hits was in fact slightly lower the respective probability of a hit following 1, 2, or 3 misses:

   $P(\text{hit}|1\,\text{miss}) = .54$ $\quad P(\text{hit}|1\,\text{hit}) = .51$

   $P(\text{hit}|2\,\text{misses}) = .53$ $\quad P(\text{hit}|2\,\text{hits}) = .50$

   $P(\text{hit}|3\,\text{misses}) = .56$ $\quad P(\text{hit}|3\,\text{hits}) = .46$

Comment on the notation:

The symbol $P(X|Y)$ denotes the probability of an event *X*, given that event *Y* is present.

   (ii) The number and length of the observed »runs«, i.e. the series of events of the same type (hits or misses), are not significantly different from that resulting from a random process with independent trials.

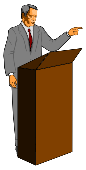
3. There were no »hot« und »cold« nights, i.e. nights with players having a »hot« hand vs. nights without players having a »hot« hand.
4. The number of free shots also revealed no indication of a »hot« hand, i.e. following a hit the probability of a hit was not higher as after a miss.
5. An experiment comprising 14 male and 12 female students of the Cornell basketball team performed shooting hoops revealed no indication of sequential dependency between hits and misses.
6. The students were also unable to correctly predict their own scoring for the next week or those of their team-mates.

7. An assessment of sequences, e.g. (**X**= hit, **O** = miss)

XOXOXOOOXXOXOXOOXXXOX,

by the 100 students from the Stanford and Cornell University (cf. Point 1) revealed that sequences were judged as randomly generated (and not as »streak« sequences, i.e. sequences containing »runs« that were generated by »hot« and »cold hands, respectively) only in case of shift probabilities being greater or equal to .70.

In conclusion, the results of the investigation of Gilovich, Vallone & Tversky (1985) reveal that the »hot hand« in basketball simply does not exist despite the strong belief of many professionals.
The fact that sequences with a shift probability of .50 were predominantly judged as »streak« sequences is an indication of the effect of the law of small numbers. Specifically it reflects the erroneous belief that each subsequence of a random sequence has to represent the random character of the process by which the sequence was generated.



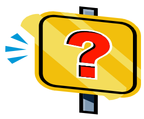
Comment 2-3:

1. The absence of the »hot hand« in basketball has been called into question (Larkey, Smith, & Kadane, 1989).
   However, the statistical facts of the independence of hits and misses on previous results cannot be doubted.

2. There seems to exist a »hot hand« in other disciplines, like golf, billiard, throwing darts, or horseshoe throwing.

2.3.2 The »Hot Hand« and the Gambler's Fallacy

Gillovich, Vallone & Tversky (1985) explained the »hot hand« by means of representativeness: Short series of hits are interpreted as indication of the »hot hand« despite missing dependencies and despite the fact that the series may well stem from a random process with independent trials.
On the other hand the gambler's fallacy is explained by peoples' belief in a balancing effect of randomness, i.e. after a long »run« the probability of the event making up the run should decrease. This raises the following issue:



Question:
Why are people believing in an increasing probability of an event of the same type (hit vs. miss) in case of basketball whereas they are believing in a reduction of the probability of an event of the same type in the case of gambler's fallacy?

The answer to this question is again found in the presence of a subjective theory: In case of basketball there exists a subjective theory that makes the »hot hand« plausible: The scoring of a player depends not only on randomness but is well dependent on the player's capabilities. The hot hand may be conceived as some sort of »priming« of the player's capabilities.

In case of tossing a coin, throwing a dice, or playing roulette, personal abilities are not relevant. Thus the subjective theory concerning the balancing effect of the randomness is invoked.

2.3.3 Rationality, the »Hot Hand«, and Gambler's Fallacy

Question:
Can the believe in the »hot hand« as well as the gambler's fallacy be regarded as rational behavior?

Ayton und Fischer (2004) claimed that the belief in the »hot hand« may be considered as instrumentally rational since it may result in a strategy that maximizes the scores of a team. For example, it might be useful to pass the ball predominantly to players with a great number of »streaks« (sequences of hits).

This line of reasoning has a major weakness: It might probably be a good strategy to pass the ball to players with a high (average) hitting rate. However, since this is independent of the »hot hand« (i.e. whether the player has scored immediately before) a strategy based on the latter makes no sense and might in fact result in negative results. Moreover, since the »hot hand« does not exist, any strategy based on this false belief seems useless (Since Ayton and Fischer (2004) do not provide any empirical evidence about positive effects of the belief in the »hot hand« their reasoning is quite weak).

Concerning the gambler's fallacy Pinker (1997) argues that in everyday life stochastically independent events are quite rare. Rather, in daily life we are confronted with limited populations as well as sampling without replacement. In this case, it makes sense to assume that the probability of an event decreases after each occurrence of the event. For example, having already observed 100 wagons of a train it is reasonable to assume that the end of the train will be reached soon.

This argument might explain the existence of the gambler's fallacy: In man's evolutionary history individuals were mainly confronted with sequences of events where the occurrence of an event, in general, reduced its probability of appearance in the future. By consequence, people with the respective idea (or subjective theory) increased their selective value.

However, Pinker's argument does not justify the claim of the gambler's fallacy being an instance of rational behavior since in case of roulette or tossing a coin the single events are (with good approximation) independent. The probability models as those considered by Pinker are not applicable.

In the following we discuss another mechanism that is an important source of erroneous contingency and causal judgments.

2.4 Regression to the Mean

The phenomenon of the regression to the mean was discovered by Sir Francis Galton (1822-1911), a cousin of Charles Darwin. During his eugenic studies he found that extreme values of characteristics were not transferred from the parents to their children. For example, sons of very tall or very short fathers are predominantly of mean size.
First, Galton provided an erroneous interpretation of the results by assuming that children’s characteristics are determined only partly by the direct parents. A portion of the characteristic was, according to Galton’s hypothesis, determined by early ancestors.
However, subsequently he found that the phenomenon also works in the opposite direction: Farthers of very tall/small sons are predominantly of medium size. Clearly, this could not be accounted for by means of heredity.

In the following, the principle of the regression to the mean is discussed in detail. This is followed by a presentation of important judgment errors that are caused by ignoring the principle.

2.4.1 Regression to the Mean: Basic Principles

Prior to the presentation of the principle let us examine an illustrative example.

Ex. 2-5: Regression to the mean

Consider the following setup:
A group of students has to take an exam with multiple choice questions.
After the test the upper and lower 10 percent (with respect to their scores in the test) are submitted to a second test comprising questions from the same pool of questions.

Result:

- ❑ In general, students from the upper 10% will perform less well in the second test compared to the first round.
- ❑ On the other hand, students from the lower 10% will, in general, get higher scores in the second test compared to the first one.

Explication:
The test results are determined by two principle factors: (1) The knowledge of the person, and (2) chance factors.
Clearly, the student's knowledge will play a role in the second test too. However, presumable the good or bad fortune will probably change between tests.

Comment:
Clearly, the phenomenon is strongest if the student's knowledge has no impact on the results of the test, and by consequence the results is completely due to chance factors.

2.4.2 Ignoring the Regression to the Mean as a Source of Judgmental Errors

The historical example, presented above, concerning Francis Galton's explanations has already illustrated the typical error resulting from ignoring the phenomenon of the regression to the mean: People invent erroneous causal theories in order to explain the phenomenon. In the following a number of examples from everyday life are presented illustrating consequences of disregarding the phenomenon.

2.4.2.1 The Flight Instructor

Kahneman und Tversky (1973) present the following nice illustration: A flight instructor of the Israeli Air Force explained Daniel Kahneman that he does not approve of positive reinforcement, for the following reasons: Whenever he praised the performance of a candidate after a successful flight maneuver the candidate performed less well the next time. On the other hand, criticism of the bad performance of a candidate resulted in a superior performance the next time.
Apparently the instructor did not understand the phenomenon of the regression to the mean according to which an excellent performance is usually followed by a less well performance (and vice versa).

2.4.2.2 Explanations in Sports and Profession

Most athletic achievements are influenced by personal abilities as well as by chance. However it seems that the impact of chance on the final outcome is underestimated by most people (among them many experts). By consequence the significance of the regression to the mean is not well recognized.
Consider the situation where a newcomer (a single athlete or a team) exhibits an extremely high performance. However, after a while the performance deteriorates considerably. In this case numerous explanations by magazines and newspapers are presented, like »he has allowed success to go to his head«. However, these explanations overlook the phenomenon of the regression to the mean. For example, the regression to the mean explains why in the German soccer league, ex-

cept for Bayern München, soccer teams have rarely won the championship in successive years (To be more concrete: Between the years 2000 and 2010 Bayern München won the championship 6 times whereas in the other cases different teams [Bremen, Dortmund, Stuttgart and Wolfsburg] became champion). It seems that only Bayern München has the potency to become champion regularly.

Numerous explanations concerning occupational attainment are also the result of a neglect of the regression to the mean (Nisbett & Ross, 1980): People are often inclined to explain the failure of excellent students in their subsequent occupation using arguments like: »She got excellent abilities but, unfortunately, she lacks the necessary drive« or »He has become too engulfed in administrative details«, or »She simply has not received the required support of her colleagues«. This type of ad hoc explanations ignores a simple fact, namely, that there may be only a modest correlation between academic success and occupational attainment.

2.4.2.3 Fairness towards Members of Disadvantaged Groups

In the United States there existed different measures to support disadvantaged black people. One of these arrangements consisted in the admission of Black students with lower scores to colleges and universities (see, e.g. Herrnstein & Murray, 1996) [*Note:* In the meantime the Supreme Court has disabled this practice].

There exist two rival theories concerning the success of promoted Black students: According to the *hypothesis of overachievement* Black students with scores located in the upper range with respect to their own group but in the lower range with respect to the White candidates (cf. Figure 2-1) should reveal an excellent academic performance.

According to the *regression to the mean* the academic development promoted Black students should be inferior to that of their White colleagues. Specifically, the future performance τ_{b} of a Black who is located in the upper range with respect to his group should be worse than the achievement x_{b} actually observed. By contrast, members of the White group with the same score x_{w} $(x_{\mathrm{S}} = x_{\mathrm{b}})$ are assumed to exhibit a higher achievement in the future (cf. Figure 2-1).

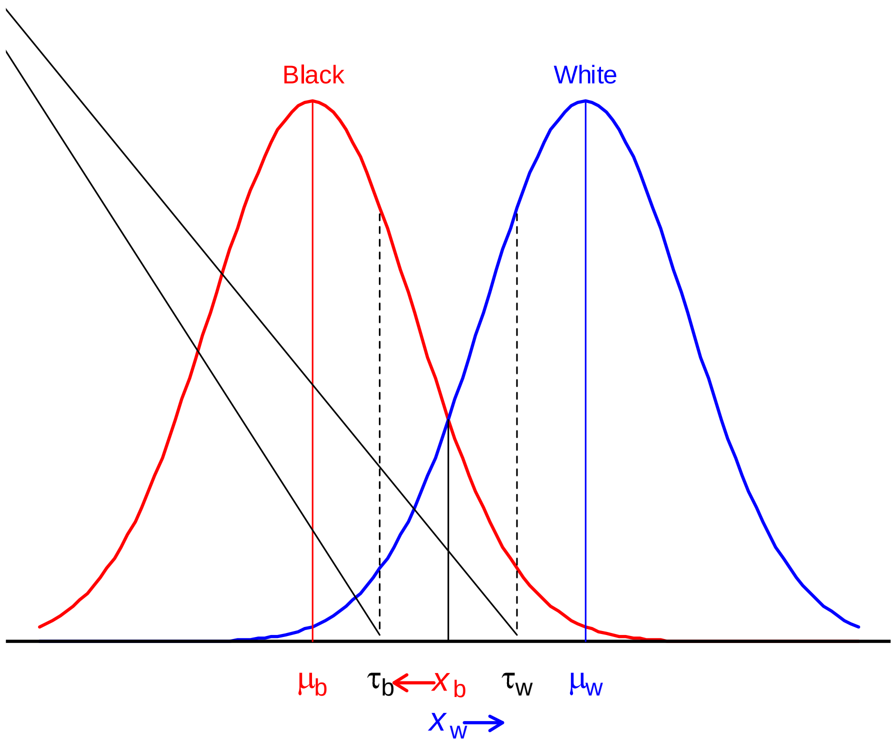


Figure 2-1: *Demonstration of the regression to the mean.*

The professor of statistics Alan Zaslavsky from the Harvard University argued that participants of a disadvantaged group, once selected, abandon the burden of their group. By consequence there will be no effect of the regression to the mean since the group of Black is no longer the relevant reference group. These hypotheses were tested in a number of studies (cf. Wainer & Brown, 2007).
In one study more than 46'000 students from 38 colleges participated. The SAT (Scolastic Aptitude Test) was used for predicting the HS-GPA (High School Grade Point Average). A clear effect of the regression to mean was observed: Students from Asia as well as White students exhibited an increase of the HS-GPA with respect to SAT whereas for Black and Hispanic students a decrease was observed.
In further study the performance of Black and White students was predicted on the basis of the scores in the introductory examination. Again, the model of the regression to the mean was confirmed: A comparisons of Black and White students with the same scores revealed that Black students are be consistently inferior in subsequent examinations.
These results are in clear opposition to the hypothesis von Zaslavsky that members of disadvantaged groups are able to abandon the bondage of their origin.
A detailed formal analysis of the regression to the mean with applications to statistics will be presented below in Section 2.6.5. We now turn to the discussion of further reasons for erroneous contingency and causal judgments.

2.5 Erroneous Weighing of Relevant Information

Numerous studies demonstrate that people ignore or do not adequately weigh information that is relevant for judging (causal) relations.

2.5.1 Assessing the Information in Contingency Tables

Contingency (cross) tables are a convenient way to represent probability information concerning the relationship of two (or more events). It is important to note that contingency tables comprise the *complete* probability information about the relationship of two events. However, untrained persons are in general unable to extract from the table the relevant information about the association of the events. The following example illustrates the issue.

Ex. 2-6: Information in contingency tables

Assume that there exists a new method of smoking cessation. In order to test the efficacy of the new method N = 340 participants were investigated.
The following criterion of success was assumed: The participant had stopped smoking for at least 30 days.

There were two groups: a treatment group that received the new treatment and a control group without any treatment.

Tab. 2-1 depicts the number of participants in the treatment and control group that were successful or not according to the given criterion.

	Success	No success
New method	200	75
No treatment	50	15

Tab. 2-1: *Number of cases with success / no success within the treatment and control group.*

Most lay persons interpret this data as revealing a positive relationship between treatment and success, i.e., they belief that the treatment has a positive effect on smoking cessation.
In fact the relation between treatment and success is slightly negative.

Cognitive Mechanism 2-4: Erroneous weighing of contingency information

Given:

A 2 × 2 contingency table (cross table) representing the joint presence / absence of two events or attributes (cf. Tab. 2-2). The four cells of table are labels by the letters (*a*)–(*d*).

	Event / Attribute 2	
Event /Attribute 1	1	0
1	(a)	(b)
0	(c)	(d)

Tab. 2-2: *2 × 2 contingency table representing the joint presence of events or the values of binary attributes: The letters (a)-(d) denote the four cells of the table.*

The information in the different cells is differentially weighted, in the following order (Schustack & Sternberg, 1981):

$$(a) > (b) > (c) > (d).$$

Thus, the information in cell (a) receives a higher weight than that in cell (b) which in turn receives a higher weight than the information in cell (c). Finally, the information in cell (d) receives the lowest weight.

Cognitive Mechanism 2-4 explains why people infer a positive association between treatment and success in Tab. 2-1: Cell (a) contains the most cases.

Method 2-1: *Assessing frequency information in 2 × 2 contingency tables:*

Given:

The 2 × 2 contingency of Tab. 2-2.

The following applies:

1. Cells (a) und (d) contain evidence about a positive association between events whereas cells (b) und (c) comprise evidence about a negative association.

   Comment:
   It is assumed that Value 1 indicates the presence of an event and 0 its absence.

2. In considering the association between the two events the four cells have to be weighted equally. Consequently, *the association between two variables can only be assessed in case of information each of the four cells being present.*
3. The following simple quantity enables a quick assessment of the nature of the association:

   $$a \cdot d - b \cdot c$$

   If this quantity is positive then the association is positive and if it is negative then the association is negative. In case of the quantity being 0 there is no association between the events.

4. The rank correlation results from the above quantity by normalizing. The resulting measure is called *Yules Q*:

$$Q = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c}.$$

Note that in the denominator is simply the nominator with the minus sign replaced by a plus sign.

Comments:

1. In case of a perfect positive association (Q = 1) cells (a) and (d) are occupied only whereas in in case of a perfect negative relationship (Q = –1) only cells (b) and (c) have entries.
2. Another important measure of the association of two binary variables in 2 × 2 contingency tables is the *odds ratio* (cf. Concept 4-5 on p. 154)

For the data in Tab. 2-1 we find:

$$Q = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c} = \frac{-750}{6750} = -0.111.$$

Consequently there exists a weak negative association between the new treatment and a success.
It is important that all cells have to be considered for the assessment of an association. It is a common error to not taking into account the information of each of the cells.

Ex. 2-7: Ignoring information in cells.

A supplement of the Swiss newspaper »Tagesanzeiger« of December 2012 contained an article of Wiebke Toebelmann titled *Was schon die Kräuterweiber wussten [What the herb women already knew]:*

> *»For example, at the University of Bern that disposes of a homeopathic research department it was found how effective naturopathic treatments can be with respect to the treatment of children with an attention deficit syndrom (ADHS). The result of the study conducted was as follows: For 80 percent of the children with this illness the clinical picture was improved by means of an individually adjusted homeopathic treatment at a rate of 50 percent.*
> *The study that had been published already in 2005 has, however, received not enough attention as complained by Klaus von Ammon, the chief of the homeopathy research«* (translated by the author).

The problem of this line of reasoning consists in missing information about the improvements in case of Placebo (without the provision of a homeopathic drug).

Without this information an assessment concerning the effectivity of the treatment is impossible.

A classical study of Hamilton und Gifford (1976) illustrates how the differential weighing of contingency information influences the perception of social groups of different sizes.

Ex. 2-8: Illusory correlation and the perception of minorities (Hamilton & Gifford, 1976)

Participants were presented 39 behavioral descriptions. Each of these 39 descriptions concerned either a member of Group A or B (In order to exclude previous knowledge the labels A and B were used to denote groups).

- ❑ 26 descriptions, 18 positive and 8 negative were assigned to members of Group A, and
- ❑ 13 descriptions, 9 positive and 4 negative were assigned to members of Group B.

In the following, Group A comprising most members is called the *majority group* and Group B be is termed the *minority group*.

Tab. 2-3 displays the number of positive and negative behavioral descriptions of the two groups.

	Behavior		
Group	Positive	Negative	Σ
A (Majority)	18	8	26
B (Minority)	9	4	13

Tab. 2-3: *Distribution of positive and negative behavioral descriptions of the two groups in the experiment of Hamilton & Gifford (1976).*

Results:

In a subsequent assessment of the two groups the Majority Group A was evaluated as more positive than the Minority Group B.

Interpretation:

Participants' evaluation can be explained by the improper weighing of positive vs. negative cases: Despite the fact that in both groups the ratio of positive and negative cases is exactly the same the number of positive cases is twice as high as high for the majority group.

Comment:

> Since the ratio of positive and negative descriptions (18/8 and 9/4) is exactly the same for both groups there is ostensibly no association between group and behavior.

The differential weighing of information may be regarded as the result of attentional effects and saliency: People pay more attention to positive cases, i.e. cases with both events being present [information in Cell (*a*)] than to other types of information. Alternatively, one might state that information in Cell (*a*) is more salient. Attentional effects and saliency, respectively, are important factor in the perception of (causal) relations.

2.5.2 Attentional Effects and Saliency

The effect of saliency is often found in the context of *attribution theory*. The latter is concerned with how people explain their own behavior as well as that of others. Furthermore, the effect of the explanations on people's own performance is investigated.
We first discuss the most influential theory of Harold Kelly. This is followed by a treatment of attribution errors and biases, respectively.

2.5.2.1 The ANOVA Model of Kelly

The ANOVA (*Analysis of Variance*) model von Harold Kelley (1967) assumes that in everyday life people behave like lay scientists by integrating data from different sources of information in a nearly optimal way.
Basically, the possible causes of behavior can be classified into three categories:

1. **Person:** The causes of a certain observed behavior are assumed to be located within the person, i.e. intention, whishes, or abilities.
2. **Object:** The stimulus or the object is assumed to be relevant factor for explaining an observed behavior (Note that under the term »object« also persons other than the acting subject may be subsumed).
3. **Situation:** The actual situation is assumed to be causally relevant for the behavior.

Ex. 2-9: Dancing performance:

Evidence to be explained:
During dancing John stepped permanently on Mary's feet.

Person based attribution:
The reasons are to be found in John's dancing abilities: He is unable to keep the rhythm during dancing.

Object based attribution:
The causes are to be found in John’s partner Mary: She permanently tries to take the lead during dancing resulting in a great deal of dissonance between the partners.

Situation based attribution:
The reasons are to be found in the actual situation: the dance floor was overcrowded. People permanently pumped into John and Mary.

The three possible explanations of Ex. 2-9 adduce different causes for explaining John’s behavior. However, this raises the general question about the possible sources used by people in making their causal judgments. Kelley assumes that people usually rely on three different types of sources of information.

1. Consensus: concerns the information about how different people behave in the same situation with the same object (or stimulus) [Information about variation of the effect with the subject].

2. Distinctness: concerns information about whether other stimuli / objects result in same of different behavior [Information about the variation of the effect with the object].

3. Consistency: concerns information about whether other situations result in different behavior for the same actor and objects [Information about variation of the effect with the situation].

Ex. 2-10: Dancing performance (continuation of Ex. 2-9*):*

Consensus information:
Consensus information informs about whether other persons stepped on Mary’s feet (during dancing).

Distinctness:
Concerns information about whether John stepped on the feet of other dancing partners.

Consistency:
This type of information provides an answer to the question of whether John had stepped on Mary’s feed at previous dancing events.

Ideally, on the basis of the information about these three types of variations the reasoning person is able to locate the cause of the evidence to be explained.

Ex. 2-11: Dancing performance (continuation of Ex. 2-9*):*

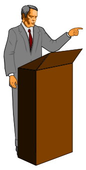
Consensus = low:
Other partners did not step on Mary’s feet.

Distinctness = low:
John also stepped on the feet of his other dancing partners.

Consistency = high:
John stepped on Mary's feet also on previous occasions.

Conclusion: The reason for John's stepping on Mary's feet is located in John's dancing performance.

According to Kelley's model the reasoning person assesses how the effect varies with the person, the object and the situation, that is, she analyses the different sources of variance. Thus one might argue that, in order to arrive at a conclusion, people perform a qualitative type of analysis of variance (ANOVA) by analyzing the three possible sources of variance.



Comment 2-4: Kelley's ANOVA model and the tools-to-theory heuristic

Kelley's ANOVA model provides a nice illustration of the tools-to theory heuristic (Gigerenzer, 1991). According to the latter, new psychological theories are the result of transforming established methods into a cognitive theory.
In the present case the statistical tool called ANOVA that is a well-established method of data analysis is assumed to be performed mentally by the reasoning persons (at least qualitatively). In this way, the ANOVA method is transformed into a cognitive mechanism.

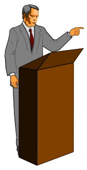
Comment:
Gigerenzer's explanation fits well with respect to Kelley's ANOVA model. However, his line of reasoning is less convincing with respect to other examples presented (like signal detection theory).
In many cases, formal theories function as starting points of theoretical reasoning. Subsequently, these models are adjusted to better explain the human data.
Decision theory provides an illustration of this fact: *Expected Utility* (*EU*) *theory* may be regarded as a starting point for modeling human decision processes. Afterwards, this account was adjusted resulting in *Subjective Expected Utility* (*SEU*) *theory*. Due to the deficiencies of this model Kahneman and Tversky (1979) developed *Prospect theory*.

Kelley's model describes the human attribution process only in very restricted sense. There are predominately two problematic aspects of the model:

1. Kelley's model is purely *data driven*, i.e. the process of causal attribution is determined solely by the given data. It completely ignores the significance of subjective theories about causal associa-

tions. However, as explicated above, the latter play an important role in causal reasoning.

2. A second shortcoming of the model consists in the neglect of saliency and attentional effects. The latter result in various biases. Specifically, on the one hand, relevant causes that are not salient receive too less weight, and, on the other hand, salient but irrelevant causes get too much attention.



Comment 2-5:

A further weakness of the model consists in the fact that motivational influences on reasoning, like self-serving biases, are ignored (cf. Section 1.2).

2.5.2.2 Ignoring Subtle Sources of Influence on the Own Behavior

Nisbett and Wilson (1977) present a number of studies demonstrating that people are incapable to recognize non-salient events that exert a pronounced influence on their behavior. The following two examples illustrate the case.

Ex. 2-12: Failure to recognize halo effects (Nisbett & Wilson, 1977):

The participants of the experiment had to assess the personal characteristics of a Belgian teacher in psychology:

- ❑ his external appearance,
- ❑ his mannerism, and
- ❑ his European accent.

The experiment comprised two conditions:

In the *Warm–Condition* the lecturer provided answers to the questions in a friendly and enthusiastic way.
In the *Cold–Condition* he answered questions in an authoriterian and intolerant way.

Results:

1. As one might expect the lecturer received higher sympathy scores in the *warm-condition* compared to the *cold-condition*.
2. A *halo effect* was found:
   Each of the three characteristics to be assessed was evaluated as attractive in the *warm-condition* whereas the same features were judged as vexing in the *cold-condition*.
3. The participants believed erroneously that the three characteristics had negatively influenced their assessment of sympathy instead of the other way round.

Interpretation:

Ostensibly, participants did not recognize the significance of the halo effect that is too subtle to be detected.

Concept 2-2: *Halo effect (Cooper, 1981; Rosenzweig, 2007):*

The *halo effect* consists in an increased correlation of judgments concerning different characteristics of an object.
The halo effect comprises different types of influences:
(i) The perception of a specific characteristic is influenced by another characteristic.
(ii) The perception of a specific characteristic is influenced by the overall impression.
(iii) The overall impression is influenced by a specific characteristic.
The term *halo effect* is used for each of these three types of influences.
The name *halo effect* is due to Thorndike (1920). He observed that during the first world war soldiers were evaluated by their supervisors either as outstanding on each of the characteristics to be assessed or as generally inferior.
The halo effect is the result of an aspiration for consistency (avoidance of cognitive dissonance).

Examples:

- Following the attack on the World Trade Center on September, 11, 2001 President George Bush was, in general, evaluated quite positively. Interestingly, also the assessment of this economic expertise improved from 47% (acceptance) to 60%. However, the latter has no relation to the events of 9/11.
  In October 2005, with the shrinkage of the acceptance of the war in Iraq and after the fiasco around the hurricane Katrina, Bush lost acceptance about *equally on each of the scales*.
- Physically attractive persons are generally judged as being more intelligent warm, sympathetic, etc.
- Scientific manuscripts from persons of well-known university are generally accepted more likely by journals.

Ex. 2-13: Failure to recognize anchor effects (Nisbett & Wilson, 1977):

Participants had to predict the behavior of a typical student of the University of Michigan in various experimental situations.

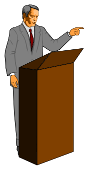
A portion of the participants received an *anchor* consisting of information about a »randomly selected student«.

Results:

1. The anchor had an ambiguous effect: From predictions biased strongly in the direction of the anchor to those tending in the opposite direction.
2. Participants exhibiting a strong anchor effect judged its influence as equally strong as those participants revealing no anchor effect.

Interpretation:

Obviously, participants did not recognize the influence of the anchor on their judgments.



Comment 2-6: Anchoring and adjustment

A cognitive heuristic called *anchoring and adjustment* will be discussed in Chapter 3.2 [cf. Cognitive Mechanism 3-5 (Page 98) and Ex. 3-4 (Page 98)].

2.5.2.3 Ignoring Consensus Information

According to Kelley's model consensus information, that is, information about the behavior of other people in the same situation (and with the same object) is used in causal judgments. However, a number of studies show that in many cases consensus information is rarely used or ignored completely.

Ex. 2-14: Ignoring consensus information
Miller, Gillen, Schenker, & Radlove (1973):

Participants were partitioned into two groups, *A* and *B*:

Group A:

Participants received the results from the experiment of Milgram (1963) revealing that nearly all subjects of the experiment had provided e-shock of substantial strength. Moreover the majority of the Milgram's subjects (65%) had applied the maximum strength.

Group B:

Participants received no information about the proportion of people who had applied the maximum strength.

Comment:

Previous to the experiment Milgram asked a number of psychiatrists to rate the percentage of subjects who would apply the maximum dose of shock. The estimates were located around 1%.

Finally, participants had to rate the values on 11 traits (like aggressiveness, warmth, grace, etc.) for two of Milgram's subjects who had applied the maximum dose of shock.

Results:

A difference in the trait ratings was found with respect to a single trait only. Thus, the information concerning the fact that most of Milgram's subjects had applied the maximum dose had no substantial influence on trait ratings that turned out as quite negative.

Interpretation:
Consensus information is base rate information, i.e. information about the prevalence of a characteristic, behavior etc. within a given population. Numerous studies demonstrate that base rate information is not adequately considered as long as there are not taken certain provisions to make it more salient (cf. Section 4.3.2).

2.5.2.4 Overestimating the Influences of irrelevant Factors

Complementary to the neglect of significant information, salient events with no influence on behavior are judged as being causally important. Here is an example.

Ex. 2-15: Erroneous assessment of reassurance on the willingness to bear e-shocks (Nisbett & Wilson, 1977):

Participants had to predict the strength of e-shocks they would accept in a subsequent investigation.

Participants of one group were provided with the reassurance that the shocks would have no negative effects on personal health. Participants of the other group received no such information.

Results:
1. The reassurance had no effect on the predictions.
2. Most participants of the reassurance group stated that the reassurance had an effect on their prediction (resulting in a higher prediction).
3. Participants of the no-reassurance group indicated that a reassurance would have increased their prediction.

Interpretation:
It might be argued that the erroneous judgment concerning the effect of reassurance is predominately due to an existing subjective (plausible) theory and less due to saliency.
However, the saliency of events is clearly influenced by subjective theories.

The examples presented reveal that, on the one hand, people are unable to recognize relevant causes of their behavior, and, on the other hand, they overestimate the effect of irrelevant factors. By consequence, people seem not to be very accurate in explaining their own behavior (or that of other people). Therefore, questionnaires should avoid question concerning causes of behavior.

2.5.2.5 The Fundamental Attribution Error

The fundamental attribution error constitutes another important phenomenon that may be explained, at least partly, by means of saliency.

Concept 2-3: *Fundamental attribution error:*

The *fundamental attribution error* consists in the tendency to attribute the behavior of a person predominately to personal factors despite the existence strong reasons in favor of situational factors.

The following two examples illustrate the phenomenon.

Ex. 2-16: Fundamental attribution error I [Jones & Harris (1967)]:

Participants received essays concerning a specific topic (e.g. an assessment of Castro's regime).

Participants either received the information that the authors of the essays were forced by the investigator to take on a certain position or they were informed that authors were completely free in their writing.

Notwithstanding the information that authors were forced to take on a certain position participants ascribed the position expressed in the essays to the personal view of the authors.

Ex. 2-17: Fundamental attribution error II [Pietromonaco & Nisbett (1982)]:

Participants were assigned to one of two groups:
Group A:
A short version of the study of Darley und Batson (1973) was presented to participants.
Group B:
A complete version of the study of Darley und Batson (1973) was presented to participants.

Digression: The study of Darley and Batson (1973) concerning helping behavior:
Participants (of the study of Darley and Batson) were seminary students who either had to give a speech about jobs or about the *good Samaritan*.
Participants were instructed that they had to give their speech in 3-5 minutes in another building where their speech would be recorded by an assistant.

The study comprised 3 conditions:

(a) Great hurry:
Participants were informed that they are late and the assistant in the other building would be waiting already. They thus should hurry.

(b) Mean hurry:
Participants were informed that it would be time to go since the assistant in the other building was probably ready.

(c) No hurry:
Participants were informed that there were still some minutes of time. However, it would be useful to already leave for the other building and to wait there. The assistant would come soon and there would be not much time to wait.

To reach the other building students had to cross a yard where Darly and Batson had placed a shabby looking ragged man who sat there with bowed head, closed eyes, and without any movement.
When a student approached the man he coughed twice and groaned without lifting his head. In case of the student offering his help the man startled and said:

> No thanks (coughing)… No, everything is ok (pause). I had just respiratory problems (coughing)... The doctor had prescribed me these pills. I have just taken one. If a stay sitting here resting a few minutes I shall be ok. … Nevertheless, thank you for your help.

Results of the study of Darley and Batson (1973):
The willingness to offer help was determined by the students' urgency. The higher the haste the lower the willingness to help.
The topic of the speech had no influence.
Result of the study of Pietromonaco and Nisbett (1982):
Participants of both groups relied on dispositional explanations: They explained the students' behavior in terms of their religious attitude.

The fundamental attribution error can be explained, at least in part, by means of saliency: The judging persons focus their attention predominantly on the acting subject and less on the situation (See also, Ex. 2-14 on p. 37 concerning the neglect of consensus information).
The differential weighing of person and situational information seems to have (also) cultural roots. Morris and Peng (1994) found differences in attributional style between US Americans and people from China. The latter are less prone to dispositional explanations and provide more situational ones compared to Americans [There was no difference between groups with respect to physical explanations].

2.5.2.6 Asymmetry of Attributions

A phenomenon that is based on differential saliency concerns the asymmetry in the explanation of behavior by the actor itself in contrast to the observers: External observers are inclined to employ dispositional explanations, i.e. the behavior of the actor is attributed to personal traits of the actor. By contrast, the actor himself tends to use a situational explanation.

For example, persons explain their aggressive behavior as having been provoked by the opposite side. Observers, on the other hand, attribute the same behavior to an increased aggressiveness of the actor.

This asymmetry of explanation is based on differential saliency: For the actor the opposite and the situation is the salient factor whereas for the external observer it is the acting person who is salient.

However, there is an alternative explanation of the asymmetry by means of *self-serving bias*: The actor might excuse her own behavior by emphasizing situational factors as the reasons for her behavior.

The attributional asymmetry can however be observed in cases where self-serving may be excluded as a possible explanation.

Ex. 2-18: Attributional asymmetry (Ross, Amabile & Steinmetz, 1977):

Participants were assigned randomly to one of two groups.

Participants of Group A had to find 10 general knowledge questions that were too difficult and specific, respectively, to be answered correctly by the participants of the other group (e.g. »Who killed the Roman emperor Phocas«? [Heraclius]; »Which is the world's largest glacier«? [Lambert glacier in Antarctica]; »What does the acronym LASER stand for«? [Light Amplification by Stimulated Emission of Radiation]).

Following each answer of Group B Group A provided feedback about the correctness, and, in case of a false answer, the correct one was presented.

Independent observers as well as the participants themselves assessed the general knowledge of participants of both groups leading to the following result:

- ❑ Participants of Group A (whose participants posed the questions) were assessed as better educated than those of Group B, by the external observers as well as by the members of Group B.
- ❑ Participants of Group A did not assess themselves as more educated than their colleagues of Group B.

Comment:

In the present case it is difficult to explain the attributional asymmetry with reference to a self-serving bias.

Let us summarize the previous discussion concerning the significance of saliency and attentional effects in the attributional process:

1. People are frequently unable to discern the real causes of their own behavior as well as that of other people.
2. One reasons for erroneous attributions are found in the fact that salient but irrelevant information / events are overestimated whereas non-salient but significant information is neglected.
3. Saliency has thus a similar effect as erroneous subjective theories. They latter may also, in part, explain differential saliency.
4. In daily life the issue of erroneous explanation of behavior gets worse due to the fact that there is rarely the time to contemplate about the reasons of one’s behavior (see, e.g. Bargh, 1994, 1997).

2.5.2.7 Criticism of Explanations by means of Saliency and Attentional Effects

In the previous section the effect of saliency on the process of attribution was demonstrated. Specifically, it was demonstrated that frequently people’s explanations are based on irrelevant but salient information whereas significant but non-salient events or information are neglected.

Explanations of human attributional errors by means of saliency, though useful, are however not fully satisfactory since (differential) saliency has to be regarded as a phenomenon that requires itself explanation: Why are specific events more salient than others?

We saw that in case of the fundamental attribution error the perceived saliency of the actor, is in part, culturally mediated: In Western cultures great importance is attached to individuality and personal responsibility. Consequently, there exists an inclination to put stronger weight on personal traits than on situational factors in the explanation of human behavior. By contrast, for members of cultures ascribing less importance to individuality the saliency of personal traits is lower.

Other types of saliency seem to have evolutionary roots. For example, in practically each culture women attach less significance to physical attraction then men.

In summary, explanations of attributional behavior by means of concepts like similarity, saliency or attentional effects may be conceived as only partial explanations since these concepts require itself an explanation.

2.6 Methodical Issues: Classical Errors and Paradoxes in Judgments of Contingency and Causality

In the present section we are dealing with errors and paradoxes that frequently appear in judgment about (causal) relationships. Contrary to the previous sections methodological issues concerning the interpretation of statistical results are the main topic of the present discussion.

The results and methods presented should be part of the basic knowledge of every student with a social science background.
Our discussion starts with an explication of the significance of the (stochastic) dependence and independence, respectively, of (random) variables. This is followed by a discussion of the problem of the inference of causal relations on the basis of associations (correlations) between variables. The next topic concerns *Simpson's paradox*. It follows a formal treatment of the regression to the mean and *Lord's paradox*. Finally, we discuss the ecological fallacy demonstrating that the latter is but a different form of Simpson's paradox.

2.6.1 On the relevance of the dependence and independence between variables

If two variables are associated the knowledge of the value on one variable reveals information about the value of the other one. The information about associations between variables is significant in at least three respects:

1. *Diagnosis:* The measurement on one variable enables conclusions about another variable that cannot be observed directly. Prominent examples are psychological tests for measuring mental constructs like intelligence, personality factors, social intelligence etc.
2. *Prediction:* Relations between quantities permit the prediction concerning quantities measured in the future. For example, knowledge about success in school may provide information about future academic performance.
3. *Causal explanation:*
   Associations between variables provide indications about the presence of causal relations. These may be used for explaining specific events. For example, The relationship between certain traits and occupational success may be causally interpreted: Certain personal characteristics, like extraversion, conscientiousness, and agreeableness may be influential with respect to occupational success. Thus, peoples' occupational success may be explained with reference to these personality traits.

Within a scientific context information about the absence of associations is of great importance since this enables one to investigate variables in isolation. The fact that not every variable is associated with any other permits the decomposition of complex systems into autonomous sub-systems that may be studied in isolation [Clearly, these sub-systems may be connected to other sub-systems by means of specific interfaces]. As an example, consider the exploration of the (human) brain. The investigation is based on the assumption that the brain consists of specialized modules with specific processing capacities.

2.6.2 The problem of inferring causal relations from observed associations between variables

(Nearly) every student in psychology (as well as other branches of the social sciences) has been confronted with a variant of the following (perfectly correct) statement:

One cannot infer causal relations from correlations.

This raises the question of why this type of inference is not justified. A straightforward reason may be found in the fact that an association between two quantities *X* and *Y* gives no indication about the direction of causation: $X \rightarrow Y$ or $X \leftarrow Y$? Frequently the causal direction may be inferred by means of substantial considerations or due to the temporal sequence of the events.

Figure 2-2 depicts the two most significant cases preventing a unique identification of the causal direction on the basis of correlational information. The rectangles, labeled *X* and *Y*, symbolize observed variables. The circles denoted by *Z* represent latent variables that are not observed (By consequence, their actual value is unknown). The arrows represent causal influences with the dashed arrows indicating possible causal relations. The arc with the double arrows represents a so called covariance arc, that is, it symbolizes the presence of a covariance (or correlation) between two variables that has not been subjected to a causal analysis.

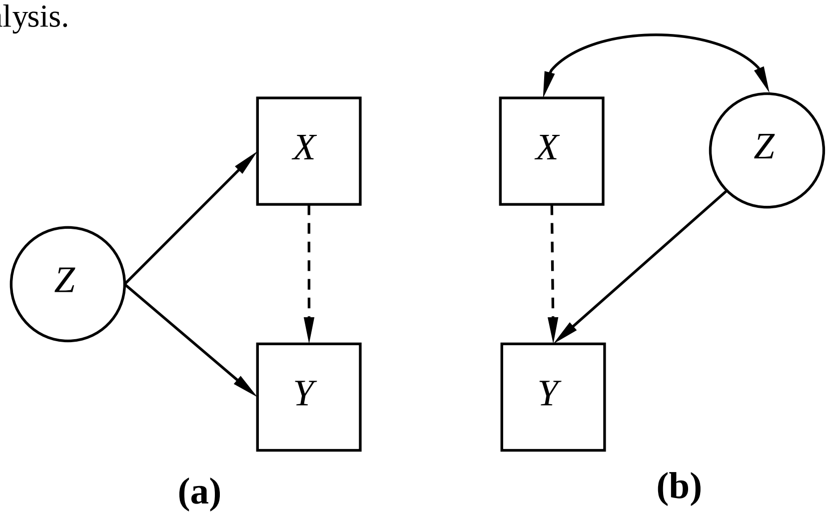


Figure 2-2: *Problems of causal interpretations of relations between variables: (a) Spurious effect: The observed association between X and Y is due to an unobserved common cause (Z); (b) Confounding: The observed association between X and Y is due to an unobserved variable Z that exerts a causal influence on Y and is correlated with the observed variable X.*

In Figure 2-2 (a) the latent variable *Z* exerts a causal influence on the observed variables *X* and *Y* thus inducing a correlation between *X* and *Y*. The dashed arrow indicates that there may be, in addition, a direct causal influence of *X* on *Y*. A situation in which a third variable has a causal effect on two other variables resulting in a correlation between the latter two is called a *spurious (causal) effect* between *X* and *Y*.

Ex. 2-19: Spurious causal effect:

In Burgenland (Austria) after World War II a correlation between the number of births and the number of storks was observed.

According to our present knowledge there does not exist any causal relationship between storks and the birth rate.

The observed correlation can be explained by the fact that after the second world war the rate of family formations increased resulting, on the one hand, in a growth of the birth rate, and, on the other hand, in an increase in the number of newly built houses, the latter serving as nesting places for storks. (*Figure* 2-3).

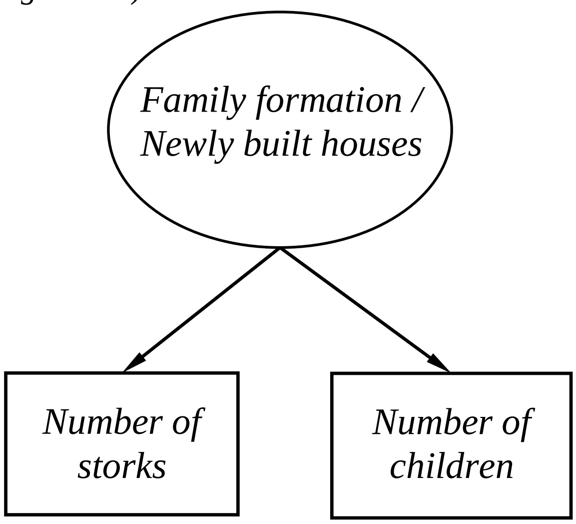


Figure 2-3: *Explanation of the association between birth rate and number of storks.*

In Figure 2-2 (b) the latent variable *Z* exerts a causal influence on the observed variable *Y*. In addition, *Z* is correlated with variable *X*. The dashed arrow indicates the possibility that there may be, in addition, a direct causal influence of *X* on *Y*.

Concept 2-4: *Confounding (causal) variable:*

A *confounding (causal) variable Z* is a variable that exerts a causal influence on an outcome variable *Y*. Moreover, Z is correlated (or more generally, associated) with the target variable *X* whose causal effect on *Y* is investigated.

In a regression context a confounding variable *Z* consists in an independent variable that is correlated with the target independent variable *X*. In addition, the regression coefficient $\beta_{Y.Z}$ of the regression of *Y* on *Z* is not zero.

Ex. 2-20: Full moon and traffic accidents:

A meta-analysis by Rotton & Kelly (1985) concerning possible effects of full moon concludes that there does not exist any relevant effect of full moon on human behavior. In contrast to common usage, they recommend at the end of the paper to abandon all subsequent investigations concerning the effect of full moon (which was no met) [However recent results arrive at the same conclusions concerning the ineffectiveness of full moon on human behavior (cf. Lilienfeld, Lynn, Rusco, & Beyerstein, 2010)].

One study found an effect of full moon on the number of car accidents. This effect was due to fact that phases of full moon were predominately observed on weekends. However the number of car accidents is clearly higher for days of the weekend due to the higher rate of circulation. Controlling for the weekly cycle makes the effect disappear.

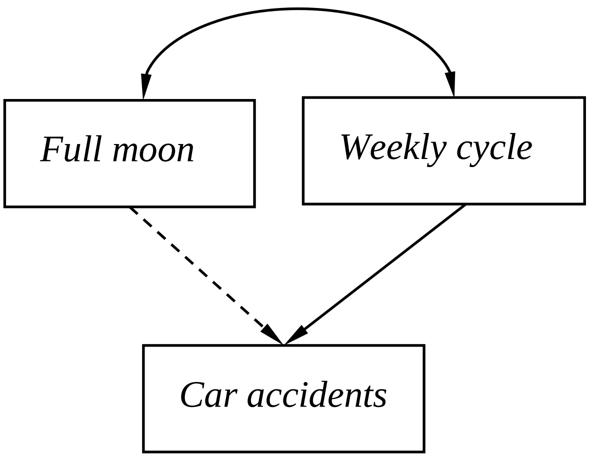


Figure 2-4: *The confounding of full moon and weekly cycle results in an erroneous inference concerning the effect of full moon on car accidents.*

The idea of one or more common latent causal variables influencing two or more other variables lies at heart of the *factor-analytic model*.

This idea dates back to Charles Spearman (1863-1945) who explained the correlation between different measures of intelligence by means of a latent common factor *g* [general intelligence] (Figure 2-5).

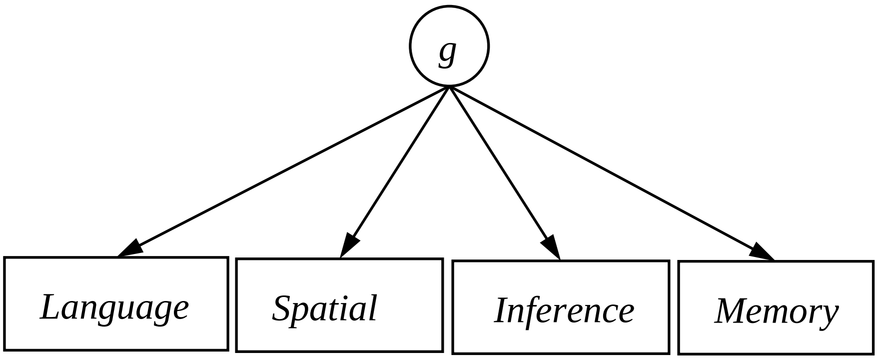


Figure 2-5: *Explanation of the correlation of various mental tests by reference to a common factor g.*

In order to establish the existence of a causal relationship between two variables X and Y and to measure the strength of the causal influence in an unbiased way the effects of possible confounders have to be controlled.
The pervious discussion about common causes and possible confounders illustrates the difficulty of establishing the existence of a causal relationship between variables beyond doubt. However there exists an important principle [well-known to most students in psychology] that is helpful:

Principle 2-1: *Experimentation and Causality:*

Causal effects can be established only by means of experiments that enable a random assignment of experimental units to different experimental conditions as well as the application of other means of control like balancing.

Let us explore the significance of this principle more closely.

2.6.2.1 Elimination of Confounding by means of Randomization and Balancing

Consider the following situation: Let the independent variable be a new treatment whose effectiveness has to be assessed. The golden standard for testing the causal efficacy of treatments consists in the, so called, *double-blind placebo-controlled studies with random assignment*. These studies comprise at least two experimental conditions, a treatment and a placebo condition [usually a control condition receiving no treatment and no placebo is also included]. The following two conditions have to be fulfilled:

(a) The assignment of clients to treatments is random.
(b) Neither the client nor the physician knows whether a placebo or a treatment is applied.

Figure 2-6 illustrates the significance of these conditions.
The ignorance of both the client and the physician about the treatment condition eliminates the effect of possible common causes, the expectations of the physician and the client. These expectations might well influence the treatment as well as the disease. Controlling for the expectations of the physician cuts the connection between the expectation and the treatment variable.
Similarly, the random assignment of clients to treatments removes (with great probability) possible correlations between the treatment and personal characteristics.
The random assignment destroys possible correlations between the treatment and confounders since the random assignment results, with high probability, in a uniform distribution of different values of confounders over treatment conditions.
As soon as the relevant connections are destroyed (cf. Figure 2-6, the red double lines) the causal effect *Treatment* → *Disease* can be estimated without bias *despite the fact that the other variables exert a causal influence on the outcome variable.*

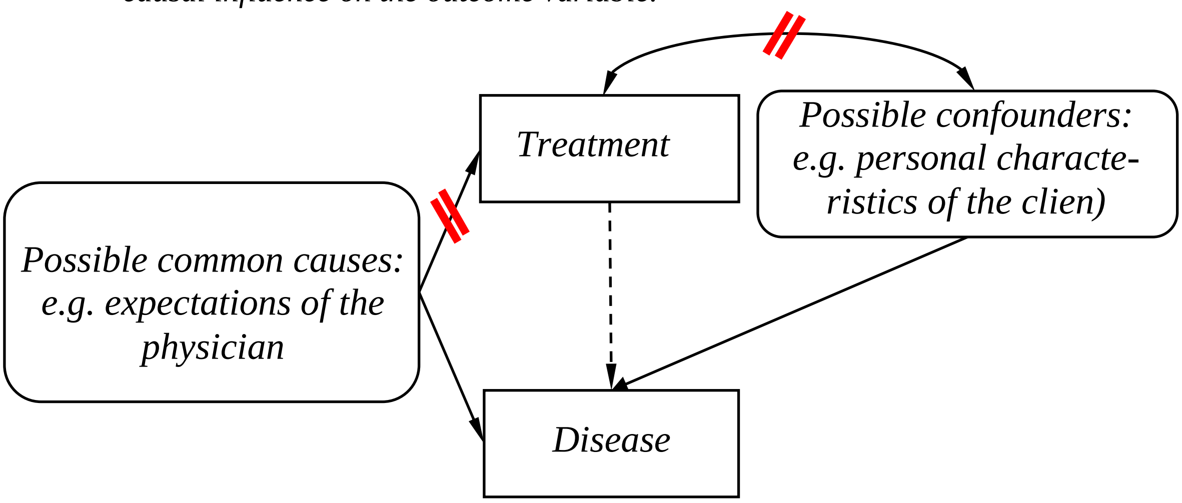
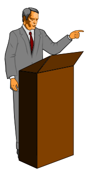


Figure 2-6: *Effects of randomization and controlling for experimenter effects: The red double lines symbolize the disruption of connections.*



Comment 2-7: Varieties of causal modeling

The model of Figure 2-6 was used for illustrative purpose only. It might be argued that the expectation of a physician may be better regarded as a factor interacting with the effect of the treatment. By consequence the variable expectation should be modeled as a confounder.

A comparison of experimental with pure observational studies (and quasi-experimental studies, respectively) reveals that for the latter correlations between confounding variables and the target causal variables cannot be removed since a random assignment of units to treatment is impossible. The possibility of statistical control requires the inclusion of covariates. However, as illustrated below, statistical control using linear regression methods is far from being optimal (cf. Section 2.6.5).

The random assignment of clients to treatments prevents with great probability the existence of (substantial) correlations between confounders and target causes. However, with small samples the values of the confounders might not be distributed perfectly uniformly over treatment levels. By consequence, for variables that are known to have an effect on the outcome variable and/or interact with the target cause the levels of the variables should be distributed uniformly over treatment levels by design. This method is called *balancing* (*parallelization*).

Having discussed the benefits of the experimental methods (compared to observational studies and quasi-experimental designs) there remains the following issue.

Question:
Is it possible to assess the causal influence of a target cause on an outcome variable without doubt (e.g. by using sophisticated experimental methods)?

The answer to this question is plainly »no«. There are, in general, two reasons for this negative verdict:

1. As mentioned above, the random assignment of units to conditions can remove possible correlations between confounders and target independent variables with great probability. Consequently, there remains a small risk that not all relevant correlations have been eliminated.
2. It is not possible to exclude with absolute certainty all hidden causal variables (i.e. variables that have not yet been identified as being causally relevant) that vary systematically with the target causal variable.

Ex. 2-21: The classical experiment of Rock (1957):

(i) Introduction:

The objective of the experiment consisted in providing empirical support that in pair association learning associations between items are not learned in an incremental way. Instead they are learned in an all-or-none fashion.

(ii) The basic problem:

There are two possibilities how associations between items of a pair in pair association learning develop:

(a) Associations are strengthened in an incremental way, i.e. in each trial the strength of the association between the items of the pair is slightly increased.

(b) In each trial an association between items of a pair is either created or not. The generation of an association is thus an all-or-none-process. In each trial associations are established in each trial for a part of the items.

(iii) Comment on the method of pair association learning:

In pair association learning, in each trial an item (the stimulus) is presented and the participant has to respond with the second item of the pair. Afterwards feedback (the correct response) is provided.

(iv) The method of replacing unlearned response items:

To test the two possible alternatives (incremental vs. all-or-none learning) the performance of two groups were compared:

(a) Control group:

Participants of this group were presented the items of a list until they could correctly reproduce the whole list.

(b) Experimental group:

Response items that were not reproduced correctly in given trial were replaced by another response. Thus the stimulus item was paired with a new response.

The replacement was performed even if the response had been reproduced correctly in a previous trial (other than the last one).

The replacing item was randomly drawn from a pool of response items from which also the original response items were drawn randomly.

(v) Hypothesis / Argumentation:

If the associations between items of a pair are strengthened in an incremental fashion participants of the control group should exhibit a superior performance (= prediction of the associative learning theory).

If, on the other hand, associations are developed in an all-or-none manner there should be no difference between groups (= prediction of the all-or-none hypothesis).

(vi) Results:

In both experiments (Experiment 1 used letter-digit pairs, Experiment 2 used nonsense syllables) the difference between groups were negligible.

Experiment 1:

❑ Mean number of trials to reach the criterion:

Control group:	4.55
Experimental group:	4.35

❑ Mean number of errors until attaining the criterion:

Control group:	17.9
Experimental group:	17.6

Experiment 2:

❑ Mean number of trials to reach the criterion:

Control group:	8.1
Experimental group:	8.1

❑ Mean number of errors until attaining the criterion:

Control group:	26.7
Experimental group:	29.2

(vii) Interpretation:

The results clearly favor the all-or-none hypothesis over the incremental hypothesis.

(viii) A possible confounding:

The method of eliminating unlearned items removes the more difficult ones from the learning set replacing them by items that may be easier to learn. Thus the advantage of the control group is traded off by the employment of a simpler learning set.

Up to now it is unclear whether the central result of Rock (1957) is due to this possible confounding (Kahana, 2012).

An interesting example of the limits of causal judgments on the basis of double-blind studies in evidence based medicine was provided be Benedetti (2014).

Ex. 2-22: Limits of randomized double-blind studies in evidence based medicine (Benedetti, 2014, p. 326-331):

Studies investigating the effects of hidden applications of drugs (e.g. the benzodiacepine Valium) revealed that the »gold standard« of randomized double-blind studies is not sufficient to rule out the possibility that the treatment is ineffective since the treatment may enhance the placebo effect without exerting a direct positive effect on the disease.

In conclusion it might be argued that the experimental method, though at present the best method for testing causal assumptions, does not provide a conclusive way to prevent erroneous causal inferences.

2.6.2.2 Adjustments for Confounders in Observational Studies

Randomization and balancing are typically used in the context of experimental or quasi-experimental studies. With pure observational studies these methods are frequently not applicable. In this case researchers usually resort to statistical techniques to control for confounders and mediators.

Ex. 2-23: Causes of coronary heart disease:

The Framingham heart study investigated various (risk-) factors that are relevant for coronary hear disease (cf. Kahn & Sempos, 1989). The following factors were investigated (together with many other variables):

- *Age*: Different age groups.
- *Sex*: Female vs. male.
- *Systolic blood pressure* (*SBP*).

The dependent variable was the incidence of a coronary heart disease (*CHD*): present vs. absent.

Assume that we would like to investigate whether participants with *SBP* ≥ 165 show a higher risk of getting a *Coronary Heart Disease* (*CHD*) than participants with *SBP* < 165.

The marginal table involving the two variables *CHD* and *SBP* reveals a positive association (Tab. 2-4) as evidenced by *Yule's Q* = 0.419.

Tab. 2-4: *Marginal association between systolic blood pressures SBP and coronary heart disease (CHD).*

	CHD	
SBP	present	absent
≥ 165	95	201
< 165	173	894

This result is difficult to interpret, however, since obviously *Age* and *SBP* are correlated: *Yule's Q* measuring the marginal association between *Age* (45-54 vs. 55-62) and *SBP* (*SBP* < 165 vs. *SBP* ≥ 165) is 0.223.

Similarly, *Sex* is associated with *SBP*, with relatively more women than men being in the higher *SBP* class: *Yule's Q* = 0.290.

Thus in order to assess the significance of the variable *SBP* with respect to getting a *CHD* one has to adjust for the two other variables.
A simple and obvious way to achieve this consists *stratification* or *conditioning on sub-populations*. In this case the association between *SBP* and *CHD* is computed for each sub-population consisting of the combinations of the levels of the other variables (in our case these are the different *Sex*-by-*Age* categories (Tab. 2-5).

Tab. 2-5: *Association between Systolic Blood Pressures (SBP), Sex, Age and Coronary Heart Disease (CHD).*

		$SBP \geq 165$				$SBP < 165$				Measures of association	
Sex	Age	+	–	*n*	*p*	+	–	*n*	*p*	*OR*	*Q*
Men	45-49	9	17	26	.346	36	147	183	.197	2.162	.367
	50-54	14	21	35	.400	35	131	166	.211	2.495	.428
	55-59	16	19	35	.457	36	105	141	.255	2.456	.421
	60-62	5	5	10	.500	13	34	47	.277	2.615	.447
Women	45-49	7	35	42	.167	13	176	189	.069	2.708	.461
	50-54	16	45	61	.262	15	153	168	.089	3.627	.568
	55-59	25	44	69	.362	18	120	138	.130	3.788	.582
	60-62	3	15	18	.167	7	28	35	.200	0.800	-.111

Notes:
\+ = Coronary Heart Disease (*CHD*) present;
– = Coronary Heart Disease (*CHD*) absent;
OR = odds ratio of the association between *CHD* and *SBP* in different sub-populations (cf. Concept 4-5 on p.154);
Q = *Yule's Q* of the association between *CHD* and *SBP* in different sub-populations.

Tab. 2-5 reveals the following pattern of results:

1. Except for women in the age class 60-62 there is a clear positive association between *CHD* and *SBP*. Thus, adjusting for the variables *Sex* and *Age* does not destroy the relationship between the two variables.
2. Both variables, *Sex* and *Age*, exert a moderating effect on the association between *CHD* and *SBP*:
   (i) Except for the women in the age group 60-62 the association increases slightly with *Age*, indicating a two-way *Age* × *SBP* interaction on *CHD*.

(ii) The association is in general higher for women than for men (with the given exception for the female age group 60-62), indicating a two-way *Sex* × *SBP* interaction on *CHD*.

(iii) The increase of the *SBP-CHD* association with *Age* is in general stronger for women than for men, thus indicating a three-way *Sex* × *Age* × *SBP* interaction on the dependent variable *CHD*.

In summary, with categorical variables the method of stratification enables one to adjust for confounders and moderators. The method has a major drawback however: It requires the presence of categorical or ordinal variables. Consequently, in case of inherently continuous variables, like *Systolic Blood Pressure* and *Age*, the variables have to be split into categories. This results in a loss of information.

The most common method used to control for confounders and moderators in the social science consists in statistical control by using *multiple regression*, linear in case of a continuous outcome variable and logistic in case of binary outcomes. The linear regression coefficient represents the mean change in the dependent variable if the respective independent variable is increased by one and the other variables are left unchanged. Including further independent variables into the regression equation can result in a significant change of the size if the regression coefficients of the already existing variables. This is due to the fact that the common variance between the existing and the new variables is partialed of all the independent variables.

If the partialing out is done »by hand«, i.e. the existing variables are regressed on the set of new variables and the residuals are computed, then the regression coefficients of the existing variables remain the same independently of whether the new variables are included into the regression equation or not.

In multiple regression adjustment for confounders is thus performed by partialling out the common variance from each of the independent variables. Note that after partialling the common variance of confounders from the target variables the correlation between the two sets of variables (confounders versus target variables) is zero, and, by consequence, the confounding no longer exists.

Ex. 2-24: Regression coefficients and the partialing out of common variance (Cohen, Cohen, West & Aitken, 2003):

The data set contains information about the following set of variables:

- ❒ *TIME:* Number of years of scientific practice of a scientist.
- ❒ *PUBS*: Number of publications.

- *CITS*: Number of citations.
- *FEMALE:* 1 = Female, 0 = Male.
- *SALARY:* The actual salary in US $ per year.

The first four variables are the independent variables with the last one being the dependent one.
The regression of salary on the independent variables results in the following regression coefficents:

$\hat{\beta}_{TIME} = 857.01$ ($p < .01$);

$\hat{\beta}_{PUBS} = 92.75$ ($p > .28$);

$\hat{\beta}_{CITS} = 201.93$ ($p < .01$);

$\hat{\beta}_{FEMALE} = \text{-}917.77$ ($p > .62$).

Thus increasing the number of citations by 1 results in a mean increase in salary of $ 201.93, and being a woman instead of a man decreases the salary by $917.77 (with the other variables being held constant).
Assume, that the variable *PUBS* and *CITS* are our target variables whereas *TIME* and *FEMALE* are considered as confounding variables.
The regression of salary on *PUBS* and *CITS* results in the following coefficients:

$\hat{\beta}^*_{PUBS} = 251.75$ ($p < .01$);

$\hat{\beta}^*_{CITS} = 242.30$ ($p < .01$);

As one might expect, the coefficients have changed and both are significantly different from zero.
Regressing both variables, *PUBS* and *CITS*, on the two confounders *TIME* and *FEMALE* and regressing salary on the residuals of the two variables results in the same coefficients as those for the regression with all variables included.

Comment:
The *t*-statistics and the associated *p*-values of the regression coefficients differ between the full analysis with all variables included and the analysis using only the two target variables (concerning the pattern of significances there is no difference in the present case). Consequently, all relevant variables should be included into the regression equation.

In addition, the residuals of the two target variables, *PUBS* and *CITS*, are uncorrelated with the two confounders, *TIME* and *FEMALE*.

Tab. 2-6: *Correlation between two sets of variables (resid = residual).*

	PUBS (resid)	CITS (resid)	TIME	FEMALE
PUBS (resid)	1.000	0.127	0.000	0.000
CITS (resid)	0.127	1.000	0.000	0.000
TIME	0.000	0.000	1.000	-0.210
FEMALE	0.000	0.000	-0.210	1.000

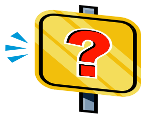
In summary, in multiple regression the adjustment for possible confounders is performed by eliminating the common variance from the independent variables. If new independent variables are added that are uncorrelated to the existing set of variables the regression coefficients of the latter remain unchanged (Their *t*-statistics and the associated *p*-values may changed, however).

2.6.2.3 The Significance of Statistical Associations For Testing Causal Structures and Models

The discussion up to now reveals that the prospects of uncovering causal relationships are rather gloomy since,

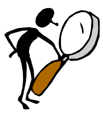
3. The presence of correlations does not enable definite conclusions about causation.
4. Even the best known method for testing causal influences does not enable one to establish causal influences without doubt.

This raises the following principle issue:



Question:
Why is it so difficult to infer causal relations?

According to my opinion the correct answer to this question was provided by the Karl Popper (1902–1994), a famous philosopher of science. He realized that the problem of inferring causal relations is but a special case of a more general problem: the problem of induction.



Concept 2-5: *The problem of induction:*

The *problem of induction* concerns the issue of whether and how inductive inferences can be justified.
In general, inductive inferences are inferences of general regularities on the basis of special cases.

The induction of causal relations can be conceived of as an instance of an inductive inference since it is concerned with the uncovering of general relations (causal relations) on the basis of concrete observations. Popper (1984) claims (a bit immodestly):

> *Selbstverständlich kann ich mich irren, aber ich glaube, ein sehr wichtiges philosophisches Problem gelöst zu haben: das Problem der Induktion [Per-*

haps I am erring but I think I have solved an important philosophical problem: the problem of induction] (Popper, 1984, p. 1).

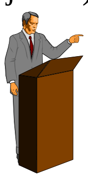
His solution is quite simple and elegant:

1. Inductive inferences cannot be justified.
2. It is however possible to rigorously test statements about general regularities on the basis of empirical data. If the statements stand these tests they are accepted provisionally. However it is always possible to refute the statements on the basis of further results that are in opposition to the statements.
3. Specifically, if there are different theories making different statements one can test which of them provides a better explanation of the data.

The solution, proposed by Popper, is in accordance with the methods used by social sciences to test and compare causal hypotheses:

1. On the basis of theoretical considerations a causal model is derived.
2. The validity of the model is assessed by comparing its predictions with observed data.

The testing of one important class of causal models, the so called *linear structural models*, employs correlations (or, more exactly, covariances) between variables: Linear structural models predict specific configurations of correlations between variables. A comparison of the model predictions with the observed structure of correlations provides information about the adequacy of the model. By consequence, if a model correctly predicts the structure of correlations between variables, and, in addition, it is plausible and not at variance with any accepted theory the model is considered as being confirmed (or not rejected).



Comment 2-8: A »nasty« practice in (causal) modeling

Unfortunately, there exists a dubious practice that seems to be quite common in applied research: Instead of creating models on the basis of theoretical considerations ahead of gathering data researcher often develop models on the basis of existing data.

The approach to develop models on the basis of existing data is in no way suspect. However, in the resulting publications studies are often described as being confirmatory despite their exploratory nature.

It goes without saying that this sort of practice is but a type of deception of the scientific community. In case of (causal) models being developed using existing data they have to be evaluated on the basis of new data in order to avoid the problem of »capitalization on chance« (i.e., the problem of modeling randomness thus thwarting generalizability).

2.6.2.4 Causal Reasoning and the Problem of Equivalent Causal Models

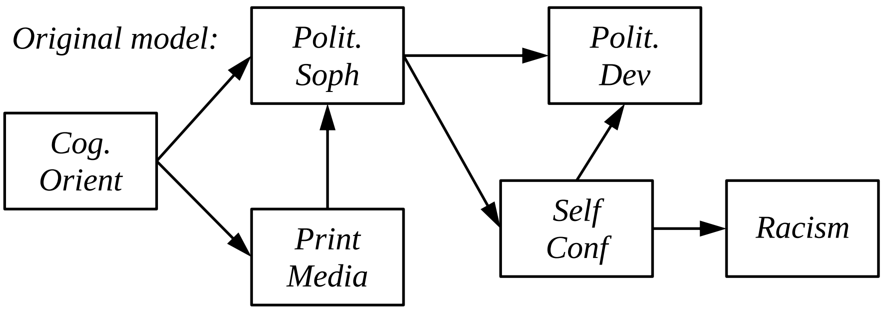
A causal models is of scientific value only if it meets the requirement that there do not exist alternative causal models that explain the data as well as the target model (or even better). This requires that the researcher considers all possible models that can explain the data equally well as the target model. The researcher has to exclude all of these models by means of substantive reasons, that is, she has to show that the models are not in accordance with existing established theories. Unfortunately, MacCallum, Wegener, Uchino, and Fabrigar (1993) have shown that in various branches, like educational science, organizational, personality, and social psychology, plausible alternative models to the existing ones exist that explain the data equally well as the one proposed by the researchers.

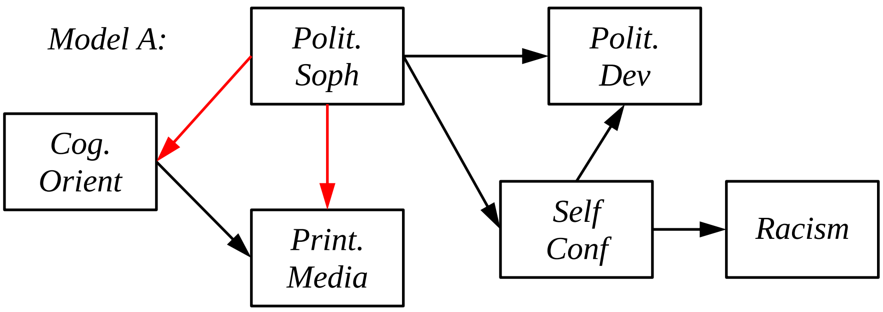
Ex. 2-25: Equivalent causal models (MacCallum, Wegener, Uchino, & Fabrigar, 1993):

Given: The following variables:

Cog. Orient: Cognitive Orientation: Interest of a person in politics.
Print Media: Print Media: Degree of print media usage.
Polit. Soph: Political sophistication: Knowledge of politics.
Polit. Dev: Political deviance: Political extremism
Self. Conf: Self confidence.
Racism: Degree of racist attidues and beliefs.

Sidanius (1988) proposed the first causal model of Figure 2-7:





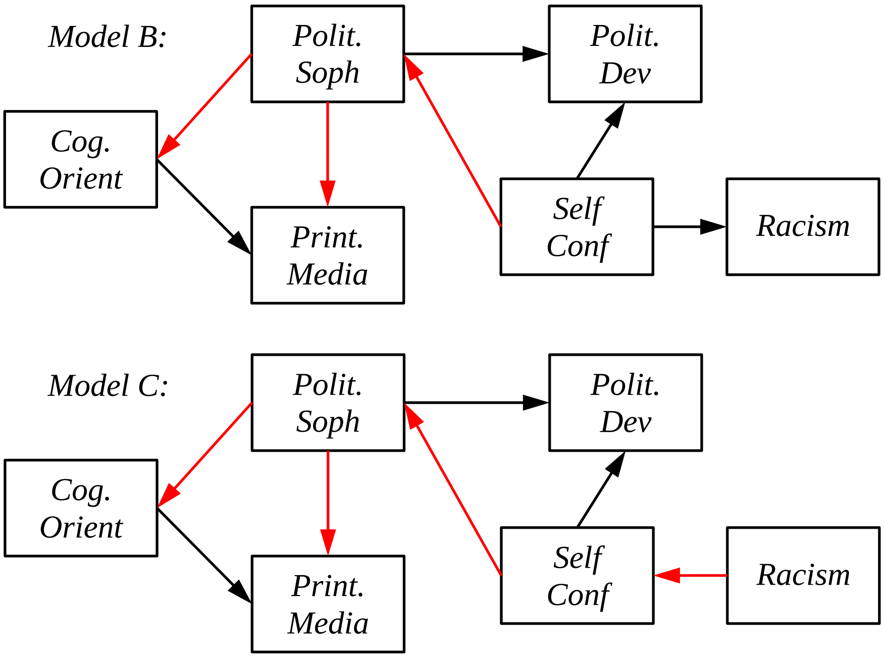
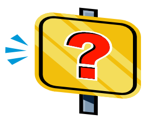


Figure 2-7: *Alternative causal models that fit the data equally well.*

> MacCallum et al. (1993) consider the alternative causal models of Figure 2-7 that explain the data equally well. Thus they cannot be differentiated from the orignal model by means of empirical data.
> In addition, according to the authors, the alternative models cannot be excluded by means of substantive reasons.

Alternative models are generated by re-orienting arrows of the original model. This raises the following issue:



Question:

Which arrows of a causal model may be re-oriented such that the resulting model is empirically equivalent to (indistinguishable from) the original model?

To answer this question, the following concept is of central importance.

Concept 2-6: *Colliders (shielded vs. unshielded:*

A *collider* is a causal structure consisting of three variables with two variables exerting an influence on the third variable. Conseqeuntly the arrows, representing causal influences, that emanate from the two variables, representing the causes, meet at the third variable.

A collider is *unshielded* if there is no direct causal connection between the two causes. Otherwise it is *shieled* (cf. Figure 2-8).

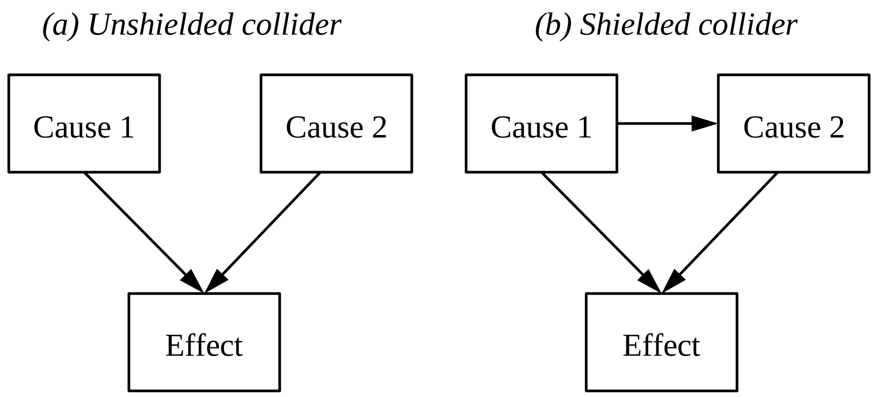


Figure 2-8: *Shielded and unshielded colliders.*

Based on the concept of an unshielded collider the following principle can be stated (that answers the above question).

Principle 2-2: *Unshielded collider rule:*

Within a causal structure arrows may be re-oriented as long as no unshielded collider is destroyed or created, and no circle is created.

Let's look at an example:

Ex. 2-26: Demonstration of the unshielded collider rule:

Figure 2-9 exhibits a causal structure with four variables: *A*, *B*, *C* and *D*. *A*, *B* and *C* exert a causal influence on variable *D*. In addition, *A* and *B* have a causal effect on *C*. Now, the arrow labed *b* may be re-oriented resulting in a new equivalent causal structure. Note that the reversion of arrow *b* destroys two colliders:

$$A \rightarrow D \leftarrow C$$
$$B \rightarrow D \leftarrow \mathrm{C}$$

In addition, the re-orientation creates two new colliders:

$$A \rightarrow C \leftarrow D$$
$$B \rightarrow C \leftarrow D$$

However, each of these colliders are shielded since in each case there is a direct causal link between the variables making up the causes. For example, in case of $A \rightarrow D \leftarrow C$ there is a direct causal link $A \rightarrow C$.

By contrast, arrows *a* and *c* must not be re-oriented since in each case the unshielded collider $A \rightarrow C \leftarrow B$ will be destroyed.

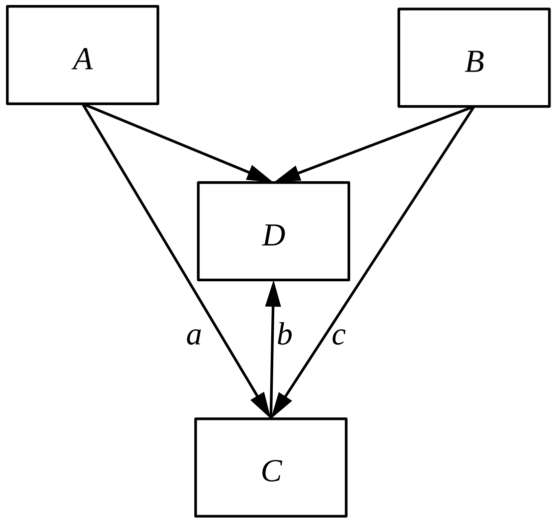


Figure 2-9: *Demonstration of the unshielded collider rule: Re-orientation of arrow b results in an equivalent causal structure, re-orientation of either a or c (or both) does not lead to an equivalent structure.*

By re-orienting causal arrows, a researcher can create new causal structures that may be either empirically equivalent to the original one or not. The unshielded collider rule enables one to check whether the re-orientation of one or more arrows results in an equivalent causal structure or not. For each equivalent causal structure the researcher has to provide arguments based on theoretical reasons of why this causal model is less convincing than that preferred by the researcher.

The final example demonstrates that a simple expansion of a causal structure by including an additional variable can transform a completely ambiguous model into an unambiguous causl model.

Ex. 2-27: The three variable mediation model and the significance of instrumental variables:

The classical example of a causal model with various statistically equivalent models is the three-variable mediator model (cf. *Figure* 2-10a). The model assumes a direct effect from the causal variable *X* on an outcome variable *Y* as well as an indirect effect of *X* on *Y* via a mediator variable *M*.

Equivalent models result by reversing arrows in the models (cf. *Figure* 2-10a-d). In fact, arrows may be oriented arbitrarily with the restriction that no closed cycle is created.

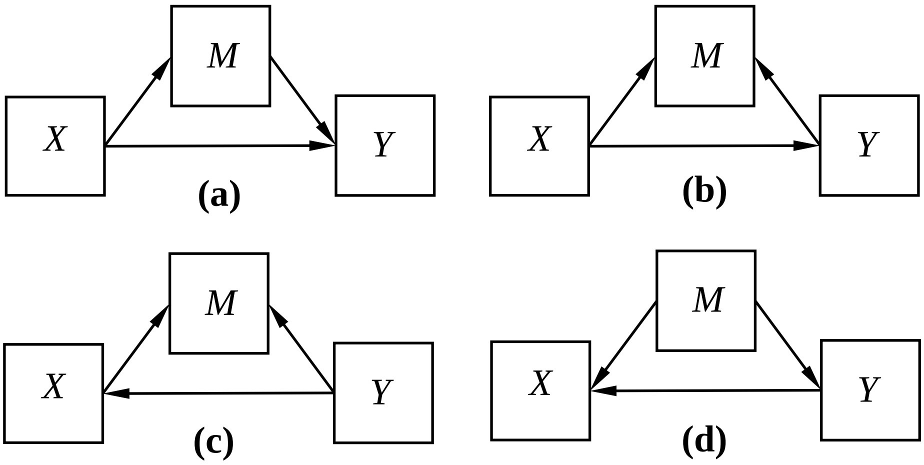


Figure 2-10: *Three-variable mediator model: (a) and statistically indistinguishable models (b) – (d) resulting from the basic model by re-orienting arrows (Only 4 of the 6 possible equivalent models are shown).*

Adding variable *I* (instrumental variable) that exerts a direct causal influence on the mediator *M* only results in a perfectly unambiguous model, i.e., no arrow can be re-oriented without changing the empirically testable predictions of the model (Figure 2-11).

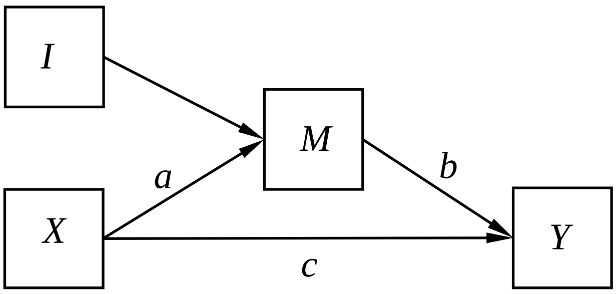


Figure 2-11: *Three-variable mediator model with an instrumental variable I that renders the model unambiguous).*

Reversing arrow *b* creates the unshilded collider $I \rightarrow M \leftarrow Y$, whereas re-orienting arrow *a* destroys the unshielded collider $I \rightarrow M \leftarrow X$. Finally, arrow *c* cannot be re-oriented since this will result in a circle. Thus the model is perfectly unambiguous.

2.6.3 Simpson's Paradox

The significance of confounding variables on the interpretation of contingency and causal relations is demonstrated drastically by *Simpson's paradox*.

***Concept 2-7:** Simpson's Paradox:*

Simpson's paradox consists in the fact that the relationship between two categorical variables can be reversed by taking an additional variable into consideration.

Here is a realistic example (New York Times Magazine, March, 11, 1979):

Ex. 2-28: Simpson's paradox: Death sentences in Florida during the years 1972-1979:

Consider the number of death sentences in the US state Florida. The issue of interest concerns the number of death sentences for Blacks and Whites. Specifically, are Blacks more often sentenced to death than Whites?
Tab. 2-7 depicts the number of death sentences for members of both ethnic groups. Apparently, there seems to be no racial prejudices against Blacks since 3.2% of White people were sentenced to death whereas the proportion of Black people was 2.4% only. Accordingly, Yules Q is negative indicating a slightly negative association between Blackness and death sentences.

***Tab. 2-7:** Death sentences in Florida for Blacks and Whites during the years 1972-1979:*

	Death sentence				
Group	Yes	No	Σ	%Yes	Yule's Q
Black	59	2448	2507	2.4	-0.16
White	72	2185	2257	3.2	
Σ	131	4633	4764		

Let us now consider an additional variable: the color of the victim. Tab. 2-8 contains the relevant data.
Surprisingly, the incorporation of the color of the victim leads to a completely different picture:

- ❑ If the victim was Black, more Black than White delinquents were sentenced to death: 0.5% vs. 0%.
- ❑ Similarly, if the victim was of White color more Black than White delinquents were sentenced to death: 16.7% vs. 3.4%.

By consequence in both parts of Tab. 2-8 (made up by the different colors of the victims) we find a high positive association between Blackness and death sentences as indicated by Yule's Q.

This raises the question how these discrepancies emerge. The following two facts are responsible for the observed divergence:

1. The killing of a Black is less often accompanied by a death sentence.
2. Blacks kill predominately Black rather than Whites. The reverse is true for Whites.

Tab. 2-8: *Death sentences for Blacks and Whites as a function of the color of the victim.*

Color		Death sentence				
Victim	Delinquent	Yes	No	Σ	%Yes	Yule's *Q*
Black	Black	11	2209	2220	0.5	1.00
	White	0	111	111	0.0	
White	Black	48	239	287	16.7	0.71
	White	72	2074	2146	3.4	
Σ		131	4633	4764		

Thus ignoring the color of the victim leads to the erroneous conclusion of a bias in favor of Blacks (i.e. Blacks would be less often sentences to death).

In fact, the contrary is the case: There is a clear bias in favor of Whites. The faulty impression results from the fact that Black delinquents kill predominately Blacks, and the killing of a Black person results less often in a death sentence.

Conclusion:

Dropping a relevant variable by summing over the different levels of this variable can results in a completely wrong judgment.

Here is another example that might well mirror a realistic scenario.

Ex. 2-29: Simpson's Paradox: Treatment effects:

A psychologist has developed a new treatment for couples. She compares her new treatment with a conventional one in two small towns: Cow-city and Goat-ville. She gets the results shown in Tab. 2-9:

The new treatment turns out as being superior to the conventional one in both towns.

Tab. 2-9: *Success of two treatmens for two cities:*

		Successful				
Locality	Treatment	Yes	No	Σ	% Success	Yule's *Q*
Goat-ville	New	20	180	200	10%	0.36
	Old	5	95	100	5%	
Cow-city	New	90	10	100	90%	0.50
	Old	150	50	200	75%	
Σ		265	335	600		

Enthusiastically our psychologist sends a report to the newspapers of both towns. The editor of the Cow-city News instructs the volunteer who had to write a short article: »Present a single table only. Otherwise readers will be confused«. Consequently the volunteer adds the data of both towns and gets the following result (Tab. 2-10):

Tab. 2-10: *Success of two treatmens summed over cities:*

	Successful				
Treatment	Yes	No	Σ	% Success	Yule's *Q*
New	110	190	300	37%	-0.30
Old	155	145	300	52%	
Σ	265	335	600		

Apparently, the new treatment is less successful than the conventional one. The volunteer writes a furious article: *The nasty statistical tricks of the Psycho lobby.*

These results are explained as follows:

1. Both treatments are more effective in Cow-city than in Goat-ville.

2. The old treatment was predominately applied in Cow-city, i.e. the environment where both treatments are more effective, whereas the new therapy was employed to a great extent in Goat-town that constitutes a more difficult environment for both treatments.

By consequence, summation over both localities results in an erroneous impression of a superiority of the old treatment since the difference in the general efficiency of both treatments as a function of location gets lost.

Here is a final example that mimics a real situation, too.

Ex. 2-30: Simpson’s paradox: Gender discrimination:

An educationalist investigates in here Bachelor thesis the study success of male and female students at the University of Freecastle for two branches: Social work and psychology. She gets the data shown in cf. Tab. 2-11.
Obviously, women are more successful in both fields.

Tab. 2-11: *Study success of men and women for two branches of study.*

Field	Sex	Success: Yes	Success: No	Σ	% Success	Yules Q
Social work	Man	127	35	162	78%	-0.20
	Woman	27	5	32	84%	
Psychology	Man	17	42	59	29%	-0.14
	Woman	92	170	262	35%	
Σ		263	252	515		

Tab. 2-12: *Study success of men and women pooled over branches of study.*

Sex	Success: Yes	Success: No	Σ	% Success	Yules Q
Man	144	77	221	65%	0.47
Woman	119	175	294	40%	
Σ	263	252	515		

The women's representative of the University would like to publish these results in ReflectUni, the journal of the University. In order to simplify the presentation she pools the results from the two branches (Tab. 2-12).

Obviously men are more successful than women. The women's representative writes a forceful article: *Discrimination of women at the University of Freecastle.*

The reason for the differences concerning the interpretation of the results of two tables is quite similar to that in the previous example: It is easier to be successful in social work than in psychology.

Moreover, women prefer psychology to social work whereas the reverse is the case for men.

Consequently, summation of the values from both partial tables (ignoring the field of study) leads to the erroneous conclusion of women being less successful than men.

The three examples presented illustrate strikingly that the removal of a variable (by summing of the values of the variable) can lead a conclusion that is in direct opposition to the one resulting from taking the whole set of variables into consideration.

Simpson's paradox raises the following question:

Question:

When is it allowed to sum over variables without changing the association between the variables in the resulting marginal table (i.e. associations are identical to those in the full table)?

Interestingly, the answer to straight forward and can be summarized in the following principle:

Principle 2-3: *Summation over variables and preservation of the table structure:*

The summation over a variable X does not change the associations between the remaining variables (in the marginal table) if X is associated with only one of the remaining variables.

In order to demonstrate Principle 2-3, let us reconsider the example concerning death sentences in Florida (Ex. 2-28 on p.63).

Ex. 2-31: Death sentences in Florida during the years 1972-1979 and association structure:

For convenience, Tab. 2-8 is reproduced below as Tab. 2-13.

Each of the 3 variables, *color of the delinquent*, *color of the victim* and *death sentence*, is associated with the other two since the color of the victim as well as the color of the delinquent is associated with the variable death sentence, and, in addition the color of the victim and the color of the delinquent are also associated.
For example, the existence of an association between the color of the delinquent and death sentence is revealed by the fact that Yule's Q is high for black victims (Q = 1.00) as well as for white victims (Q = .71) [cf. Tab. 2-13].

Tab. 2-13: *Death sentences in Florida for Blacks and Whites during the years 1972-1979:*

Color		Death sentence				
Victim	Delinquent	Yes	No	Σ	%Yes	Yule's Q
Black	Black	11	2209	2220	0.05	1.00
	White	0	111	111	0.00	
White	Black	48	239	287	16.70	0.71
	White	72	2074	2146	3.40	
Σ		131	4633	4764		

The structure of the dependencies between the three variables of Ex. 2-28 is depicted in Figure 2-12. The lines between the variables (represented by the boxes) indicate associations between pairs of variables. Figure 2-12 reveals that summation of any of the variables can result in a change of the association structure in the marginal table. This is due to the fact that each variable is associated with the other two.

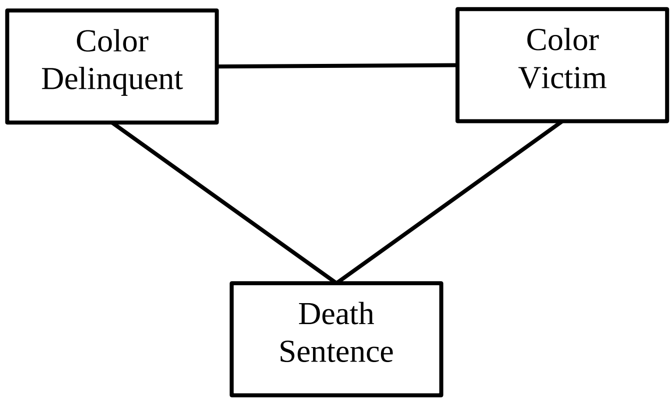


Figure 2-12: *Representation of dependences between the variables in Ex. 2-28: Lines represent dependencies between variables.*

Let us modify the data by eliminating the relationship between death sentence and color of the victim, i.e. let us assume that the risk of being executed is independent of whether the victim is Black or White. Otherwise the resulting data should be as close as possible to the original data.[1] The resulting data are shown in Tab. 2-14.

Tab. 2-14: *Death sentences in Florida for Blacks and Whites during the years 1972-1979. Data were modified to implement the independence between death sentence and color of the victim:*

Color		Death sentence				
Victim	Delinquent	Yes	No	Σ	%Yes	Yule's *Q*
Black	Black	52	2168	2220	0.024	-0.16
	White	4	107	111	0.032	
White	Black	7	280	287	0.024	-0.16
	White	68	2078	2146	0.032	
Σ		131	4633	4764		

Figure 2-13 provides a graphical depiction of the structure of dependencies.

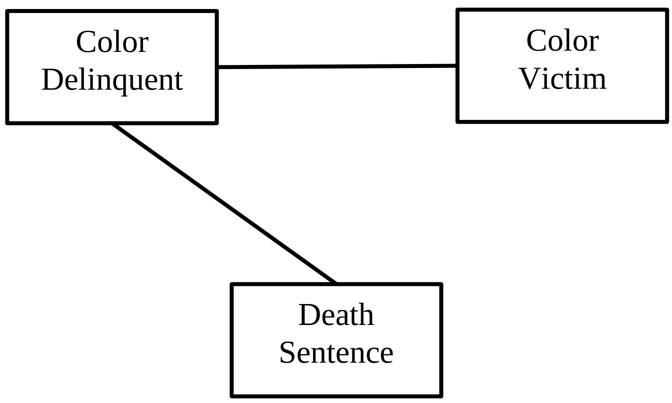


Figure 2-13: *Representation of dependences between the variables for the modified data: Lines represent dependencies between variables.*

For the data in Tab. 2-14 summation over *color of the victim* results in the same marginal table as for the original data (cf. Tab. 2-7, on p.63). The association between *color of the delinquent* and *death sentence* in

[1] The data were generated by means of a log-linear model assuming that death sentence and color of the delinquent as well as color of the delinquent and color of the victim have the same marginal frequencies as the data in the original table, whereas there is no association between death sentence and color of the victim.

the marginal table is equal to the conditional association between the *color of the delinquent* and *death sentence* for each of the levels of the third variable *color of the victim* (Yule's Q = -0.16). This is due to the fact that *color of the victim* is only associated with the *color of the delinquent* (cf. Figure 2-13).

Likewise, summation over *death sentence* leads to the same marginal table as for the original data, and, again. The marginal association is identical to the conditional associations between *color of the victim* and *color of the delinquent* for each level of *death sentence* (Yule's Q = .99). Again, this is due to the fact that *death sentence* is associated only to one of the other variables and not to both.

However, summation over the *color of the delinquent* does not result in the same association between *color of the victim* and *color of the delinquent* in the marginal table, compared to the full table (Why?).

In the following it will be demonstrated that a similar phenomenon can be observed in the context of (linear) regression where the phenomenon is denoted differently.

2.6.4 The Ecological Fallacy

The ecological fallacy represents the analogue to Simpson's paradox in the context of regression.

Concept 2-8: *Ecological fallacy:*

Given:

A hierarchical (clustered, layered) sample, i.e. the whole sample consists of subsamples from different populations.

The *ecological fallacy* results from ignoring the layered structure of the sample. By consequence, the sign of the regression coefficient computed from the whole sample (ignoring the hierarchical structure) can be different to the sign of the regression coefficients within each subsample.

Ex. 2-32 illustrates the issue.

Ex. 2-32: Ecological fallacy:

A study investigates the association between achievement and inclination to aggression in High School. The whole sample is made up of three different classes of a single school.

In Figure 2-14 the three populations (classes) are represented by the three ellipses. The symbols in grey color represent the single pupils from the three classes.

The red dashes regression line:

$$Aggression = 4.52 + 0.75 \cdot Achievement$$

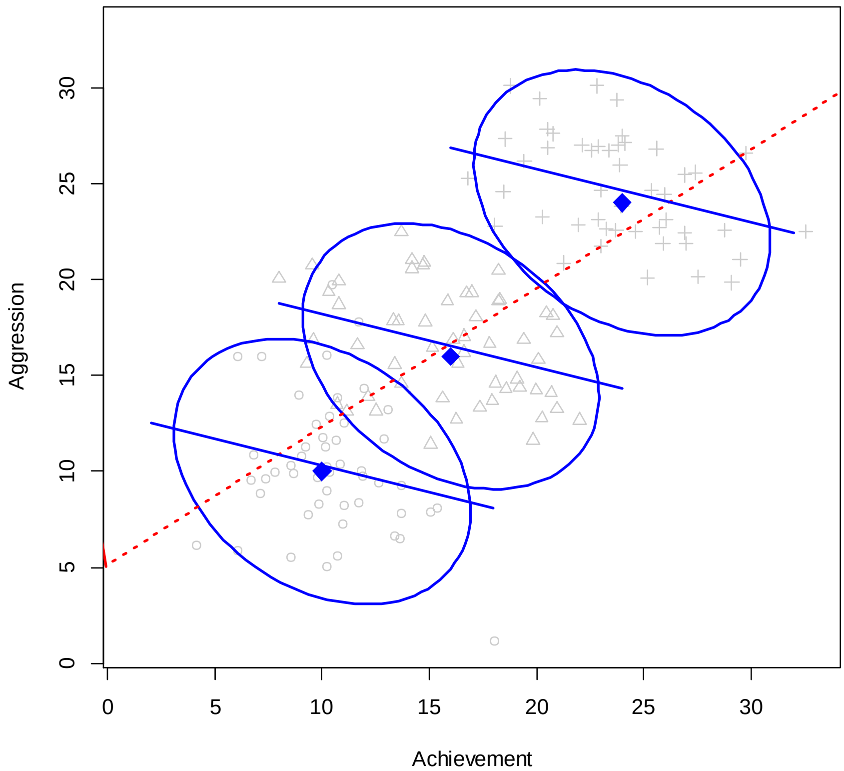


Figure 2-14: *Illustration of the ecological fallacy. The slopes of the regression lines (blue lines) are negative in each of the three groups. The slope of regression line (red dashed line) for the whole data set (ignoring the layered structure) is positive.*

leads to the erroneous impression of a highly significant positive association between *aggression* und *achievement*.
Consideration of the hierarchical structure results in a significantly negative relationship between both variables.
Interpretation:
The pattern of results is due to the fact that *aggression* increases with *achievement* on the level of the classes (= positive effect of the cluster). However, on the level of the individual pupils within classes *aggression* and *achievement* are negatively correlated.
Thus, the effect of the variable *achievement* on *aggression* is positive on the level of classes and negative on the level of pupils within the classes.

Ex. 2-32 demonstrates that, similarly to Simpson's paradox, ignoring a variable can result in an erroneous interpretation.

This concludes our discussion of erroneous judgments due to ignoring relevant variables. We next turn to our final topic: a discussion of the formal aspects of the regression to the mean.

2.6.5 Regression to the Mean and Lord's Paradox

Figure 2-15 illustrates the basic statistical aspects of the regression to the mean. The blue ellipse represents the 95% confidence region of a bivariate normal distribution with a correlation of $\rho_{xy} = 0.5$.

The ellipse consists of points of equal density on the bivariate normal density curve (cf. Figure 2-16). The 95 percent confidence region contains 95 percent of the volume under the normal density curve.

The regression to the mean is illustrated in Figure 2-15 by the fact that *the slope of the regression line (red line) is lower that the slope of the principal axis (blue line)* [in case of positive relationship between *x* and *y*.

This follows immediately from the equation for the slope of the regression line (see Figure 2-15) and the fact that the slope of the principal axis corresponds to the quotient σ_y / σ_x and the absolute value $|\rho_{xy}|$ of the correlation coefficient is always lower or equal to 1: $|\rho_{xy}| \leq 1$. Consequently the slope of the regression line is always lower than that of the principal axis except for the degenerate case where $|\rho_{xy}| = 1$. In the latter case the ellipse is shrunk to the line represented by the principle axis.

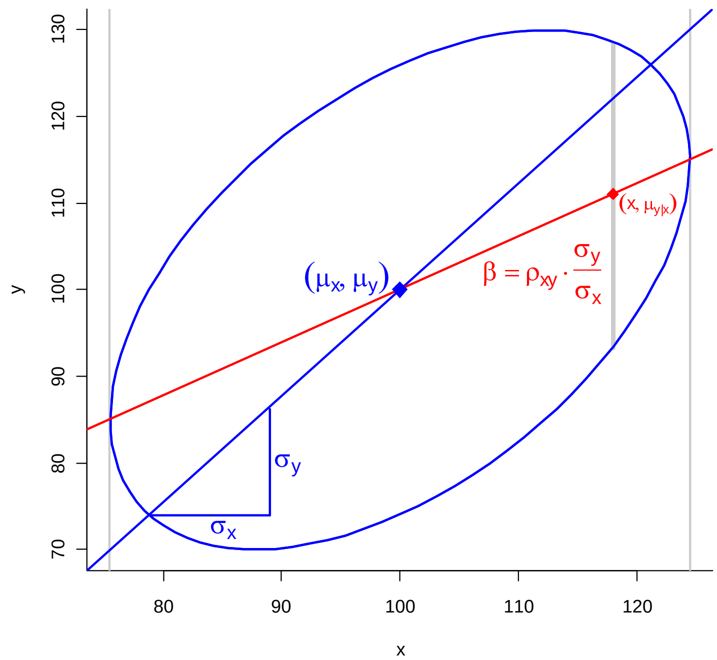


Figure 2-15: *Illustration of the regression to the mean: The slope of the regression line (red line) is lower than that of the of the principle axis (blue line). The regression line cuts the*

ellipse at the points where the vertical tangent lines (thin vertical grey lines) meet the ellipse.

Inspection of Figure 2-15 reveals two particularities:

1. The regression line cuts the ellipse and the vertical tangent lines (the thin vertical grey lines) at the same point.
2. The regression line cuts the vertical line reaching from the lower end to the upper end of the elliptic curve (the thick grey line) exactly in the middle [see the red diamond labeled (x, $\mu_{y|x}$)], i.e., each point on the regression line corresponds to the (conditional) mean $\mu_{y|x}$ of the y-values given the value x.

From Figure 2-15 it becomes immediately clear that, due to the regression to mean (i.e. the regression line is located below the main axis of the ellipsis for all values of x greater than the mean μ_x), for each selected value of x that is located above the mean μ_x the mean of the y-values $\mu_{y|x}$ for the items with the given value x will be lower than x. Thus, for each selected value $x > \mu_x$, one would expect a value of y that is lower than x since $\mu_{y|x} < x$. Similarly, for each selected value $x < \mu_x$, one would expect a value of y that is higher than x since, in this case, $\mu_{y|x} > x$. These statements are true in case of equal scales (i.e. $\sigma_x = \sigma_y$) only.

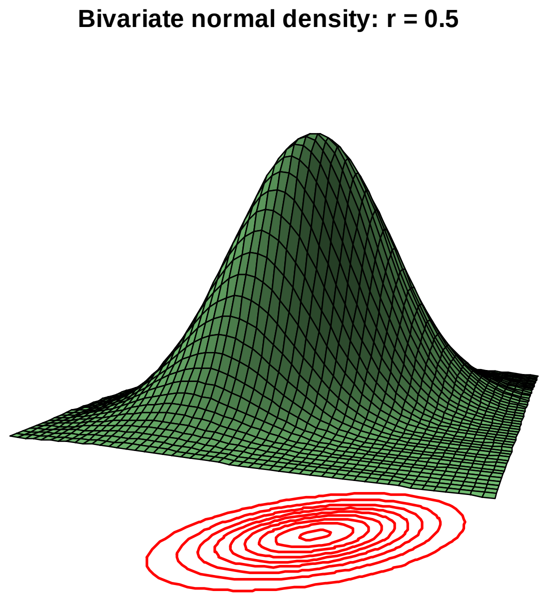


Figure 2-16: *Illustration of the bivariate normal density curve with $\sigma_x^2 = \sigma_y^2 = 1$ and $\rho_{xy} = 0.5$. The red ellipses below the*

curve represents horizontal cuts though the curve on different levels of height.

In order to illustrate the stated relationships more formally, consider the equation for predicting the value of *y* for a given *x*-value:

$$\tilde{y}_i = \mu_y + \beta \cdot (x_i - \mu_x), \qquad (2\text{-}1)$$

where the symbols have the following meaning:

$\tilde{y}_i$ is the predicted value of unit *i*.

μ_y is the mean of the dependent variable *y*.

β is the regression coefficient.

x_i is the value of unit *i* on the independent variables.

μ_x is the mean of independent variables.

Equation (2-1) and the fact that $\beta = \rho_{xy} \cdot (\sigma_y / \sigma_x)$, as well as the fact that $-1 \leq \rho_{xy} \leq 1$ result in the inequality:

$$\frac{|\tilde{y}_i - \mu_y|}{\sigma_y} \leq \frac{|x_i - \mu_x|}{\sigma_x}. \qquad (2\text{-}2)$$

This inequality tells us, that the distance between the predicted value $\tilde{y}_i$ and its mean on the *y*-scale (i.e. in units of σ_y) is always smaller or equal than the distance between *x* and its mean on the *x*-scale (i.e. in units of σ_x). If the units of both scales are equal (i.e. $\sigma_x = \sigma_y$) than Equation (2-2) amounts to the predicted is always closer to the respective mean μ_y (or, in the extreme case, equally distant) compared to the distance between *x* and its mean μ_x. This corresponds exactly to the assertion made above.

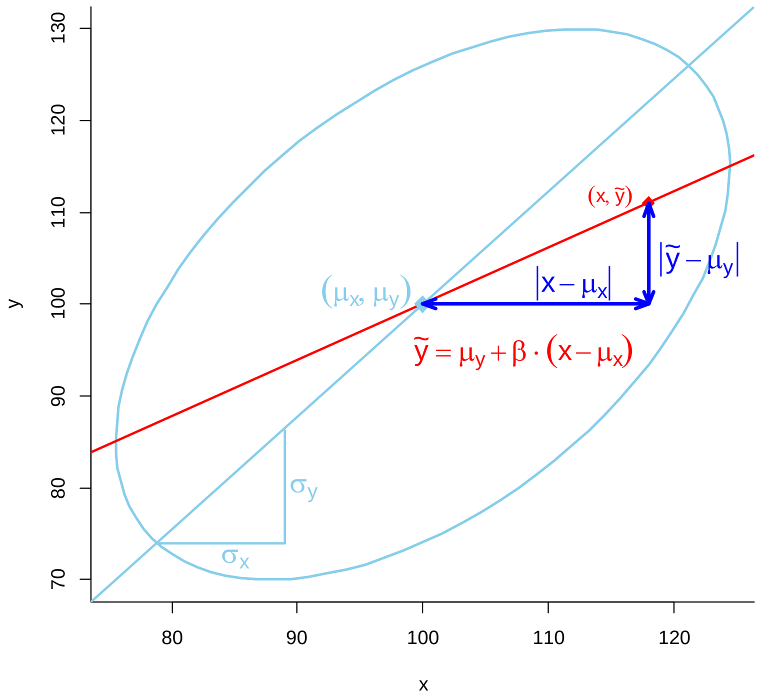
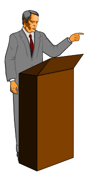


Figure 2-17: *Illustration of the regression to the mean: The slope of the regression line (red line) is lower than that of the of the principle axis (blue line). As a result, assuming equal scales, i.e.* $\sigma_x = \sigma_y$, *the inequality* $|\tilde{y}_i - \mu_y| \le |x_i - \mu_x|$ *holds.*

Figure 2-17 illustrates the inequality, assuming equal standard deviations of *x* and *y*.
The linear least square prediction corresponds to the conditional mean: $\tilde{y}_i = \mu_{y|x_i}$, i.e. the mean of the *y*-values for all items with *x*-value equal to x_i, if the conditional means $\mu_{y|x}$ are linear in *x*. In case of bivariate normal distributed data, as shown in Figure 2-15, this condition is fulfilled, and, by consequence, $\tilde{y}_i = \mu_{y|x_i}$.



Comment 2-9: On the form of the prediction equation:
The linear least square equation is usually written as:

$$y = \alpha + \beta \cdot x, \qquad (2\text{-}3)$$

with: $\alpha = \mu_y - \beta \cdot \mu_x$.

Substituting for α, gives the form of Equation (2-1):

$$y = \mu_y - \beta \cdot (x - \mu_x). \qquad (2\text{-}4)$$

From Equation (2-1) it follows immediately that in case of a zero correlation ($\rho_{xy} = 0$) and, by consequence $\beta = 0$, it is the case that

$\tilde{y}_i = \mu_y$, for all values of x. Thus, in case of uncorrelated x- and y-values the best prediction of y is the (unconditional) mean μ_y for all x-values.

The prediction equation (2-1) reveals another interesting aspect. The formula includes the mean μ_y as one component. By consequence, the same observed value of x will result in different predictions for different groups, with the predicted value for the group with a lower mean resulting in a lower prediction (cf. Figure 2-1 on Page 27).

The regression to the mean can result in an erroneous interpretation of possible effects of the independent variables in case of using an analysis of covariance to control for covariates. This phenomenon has been first reported by psychometrician Frederick Lord in 1967. It was thus termed Lord's paradox.

Concept 2-9: *Lord's paradox:*

Lord's paradox consists in the fact that an analysis of covariance for testing the difference between two groups due to an intervention, and, at the same time, controlling for initial differences between groups leads to a different result than a t-test on the differences between pre- and posttest.

Comment:

One might expect that differences between groups existing prior to the treatment are partialed out in the analysis of covariance. By consequence, initial differences should have no effect on the interpretation of differences between groups in the posttest.

This somewhat abstract characterization can be illustrated by means of a simple example.

Ex. 2-33: Lords Paradox:

Given:

Two groups of students:

- ❑ The first group switches from the secondary school to the Grammar school.
- ❑ The students of the second group stay (for the moment) in the secondary school (switching later on to the Grammar school).

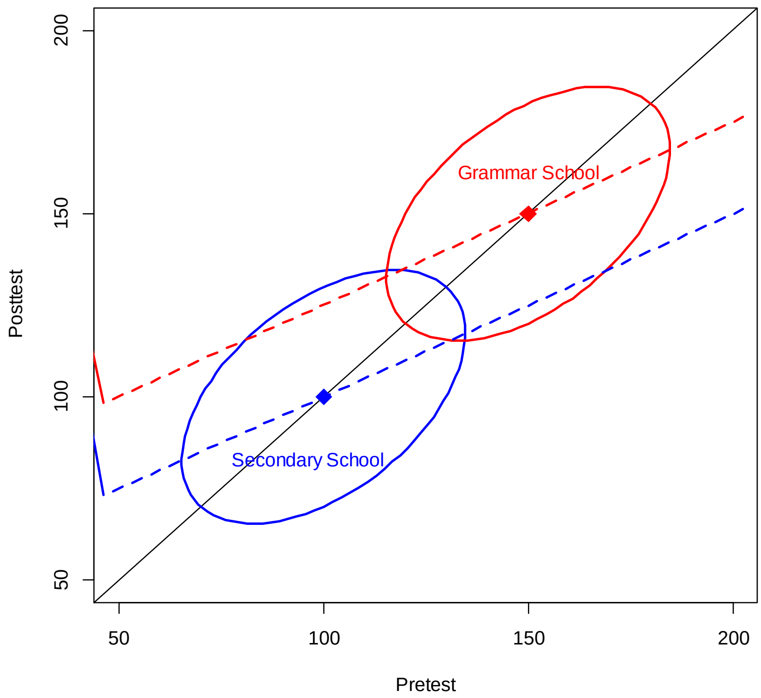
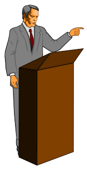


Figure 2-18: *Illustration of Lord's paradox: Despite a lack of difference between pre- and posttest measures the analysis of covariance on the posttest reveals a significant difference between groups. The difference is due to the regression to the mean resulting in different intercepts of the (dashed) regression lines for the two groups.*

Previous to the switch of the first group of students a language test is applied to both groups (=pretest).
A second test (=posttest) is applied to both groups one year after the switch of the students of the first group.
Figure 2-18 reveals that groups differ by the same amount (50 points) in the pre- and posttest. Consequently, the switch to the Grammar school did not lead to an improvement with respect to the secondary school.
As a result, a *t*-Test on the differences between pre- and posttests reveals no significant difference between groups.
On the other hand, an analysis of covariance of the posttest scores of both groups with *pretest* as a covariate reveals a significant effect of the factor *group*: Students switching to the Grammar school are superior to those staying in the secondary school (cf. Exercise 2-11).
The reason for this difference is found in the regression to the mean that leads to different intercepts of the regression lines for the two groups (cf. *Figure* 2-18).



Comment 2-10: Experimental vs. statistical control of possible confounders

Lord's paradox can be conceived of an illustration of a common wisdom:

Control by design is usually superior to statistical control.

Consequently, it is a good practice to exclude possible confounders a priori by utilizing an adequate research design instead of applying statistical tools a posteriori to control for the effects of possible confounders.

2.7 Summary

The following psychological mechanisms are found to be responsible for erroneous contingency and causal judgments:

- ❑ Erroneous personal theories concerning causal or diagnostic relationships;
- ❑ A faulty conception of the nature of random processes as well as an underestimation of the significance of randomness in daily life;
- ❑ Ignoring and faulty weighting, respectively, of relevant information in contingency tables;
- ❑ Saliency and attention effects: The influence of salient events is overestimated whereas the effect of non-salient events is ignored.
- ❑ Fundamental attribution error.
- ❑ Ignorance or lack of understanding of the regression to the mean;

The following methodological aspects were discussed that are relevant for correct judgments of contingency and causal relations. They were considered as being helpful in avoiding judgmental errors:

- ❑ A simple index, *Yule's Q*, enables a quick and convenient estimation of the strength of an association between two variables in 2 × 2 contingency tables;
- ❑ A consideration of the relevance of confounding variables as well as common causes in the evaluation of causal relationships;
- ❑ The importance of taking statistically equivalent causal structures into consideration;
- ❑ *Simpson's paradox* describes a common phenomenon. It makes clear that summing entries within a contingency table over a confounding variable (i.e. generating a marginal table that no longer contains the confounding variable) may result in an erroneous judgment. In extreme cases the judgments based on the full and on the marginal table may be in the opposite direction.
- ❑ The *ecological fallacy* may be conceived of as an analogue to Simpson's paradox in the context of regression. In this case the relevant variable ignored is usually a grouping (or cluster) variable. Similarly to Simpson's paradox the judgment taking the grouping

variable into account can result in an interpretation in opposition to the judgment ignoring the grouping variable.

- The statistical rationale of the regression to the mean can be easily illustrated graphically in the bivariate case: The slope of the regression line is always lower than that of the principal axis (except in case of a perfect correlation between the two variables). A point on the regression line represents the conditional mean of the dependent variable Y given a fixed value $X = x$ of the independent variable. In case of a multivariate normal distribution the regression estimates (conditional means) are optimal in the sense that they represent predictions with the smallest expected error.
- *Lord's paradox* constitutes a phenomenon that is due to the regression of the mean. It illustrates the limits of the analysis of covariance in controlling influences of relevant covariates.

2.8 Exercises

Exercise 2-1:

Given:

An academic achievement test that is moderately correlated with the GPA (Grade Point Average): ρ = 0.1.

Tab. 2-15: *Percentile values of the scores of an achievement test and of the GPA.*

	Scores	
Students	Achievement test	GPA
Upper 10%	>782	>3.9
Upper 20%	>685	>3.4
Upper 30%	>615	>3.1
Upper 40%	>556	>2.8
Upper 50%	>500	>2.5

Tab. 2-15 contains the percentile values for both tests. The table should be read as follows:

The 10% of the students with the highest performance in the achievement test received scores greater than 782.

The 10% of the students exhibiting the highest performance with respect to GPA achieved scores greater than 3.9.

1. A randomly selected student has got a value of 725 in the achievement test. Which value would you predict for this student with respect to the GPA. Justify your judgment.

2. Use linear regression to predict the GPA score on the basis of the observed value (725) for the achievement test assuming the following correlation coefficients:

   $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$.

   Assume that scores of both tests conform to a bivariate normal distribution.

Hints:

(i) The standard deviations of the two variables have to be derived on the basis of the percentile scores assuming normal distributions of the two measures.

(ii) The regression equation for predicting the value y of the dependent variable Y on the basis of the value x of the independent variable X is given by:

$$\hat{y} = \mu_y + \beta \cdot (x - \mu_x),$$

where:

$\hat{y}$ denotes the predicted value (the conditional mean).

μ_y symbolizes the mean of the dependent variable Y.

μ_x denotes the mean of the independent variable X.

β represents the (unstandardized) regression coefficient.

Exercise 2-2:

In 1976 the chief of the police department of Denver was dismissed for the following reasons:
Since the inception of the chief in 1971 the rate of crimes raised by 14 percent.
How should the responsible person proceed in order to better justify her decision?

Exercise 2-3:

Given: The following problem of Fong, Krantz, and Nisbett (1986):

Bert H. has a job checking the results of an X-ray scanner of pipeline welds in a pipe factory. Overall, the X-ray scanner shows that the welding machine makes a perfect weld about 80% of the time. Of 900 welds each day, usually about 680 to 740 welds are perfect. Bert has noticed that on some days, all of the first 10 welds were perfect. However, Bert has also noticed that on such days, the overall number of perfect welds is usually not much better for the day as a whole than on days when the first 10 welds show some imperfections.

Why do you suppose the number of perfect welds is usually not much better on days where the first batch of welds was perfect than on other days?

Exercise 2-4:

Please give an example of:
(a) A spurious effect;
(b) A confounding variable.

Exercise 2-5:

Given: The Framingham heart data study of Tab. 2-5 (p.53). Compute *Yule's Q* of:

(a) The marginal association between *Age* (45-54 vs. 55-62) and *Coronary heart disease* (*CHD*) [present or absent].
(b) The marginal association between *Sex* (women vs. men) and *Coronary heart disease* (*CHD*) [present or absent].
(c) The conditional association of *Age* (45-54 vs. 55-62) and *Coronary heart disease* (*CHD*) [present or absent] for the two sexes (*Sex* as stratification variable).

Exercise 2-6:

Given: The data of Cohen, Cohen, West and Aiken (2003). Compute:

(a) The linear regression of *SALARY* on *TIME*, *PUBS*, *CITS*, and *FEMALE*.
(b) The residuals of the linear regression of *PUBS* on *TIME*, *CITS*, and *FEMALE* (and the regression constant).
(c) The linear regression of *SALARY* on the residuals computed in (b).

Comment:
The regression coefficient resulting from (c) must be the same as that found in (a).

Hint:
The R function `resid` applied on the output of the R function `lm` (for perfoming linear regressions in R) returns the residuals.

Exercise 2-7:

Tab. 2-16 exhibits the efficiency of two drugs, A and B. the numbers *x*/*y* indicate that there were *x* positive cases (indicating that the drug is efficient) out of *y* cases.

	Drug	
Sex	A	B
male	5/10	38/100
female	30/100	2/10
Σ	35/110	40/110

Drug A is more efficient for male as well as for female subjects. However, when summed over both sexes Drug B is more efficient than A.
Explain the reason for the observed discrepancy.

Exercise 2-8:

Tab. 2-16 exhibits the relation between the variables body weight, weight at birth, and blood pressure (Tu, Gunnell & Gilthorpe, 2008):

Tab. 2-16: *Cross tabulation of the variables blood pressure (normal vs. high), actual body weight (≤ 90 kg vs. > 90kg) and weight at birth (low vs. high).*

		blood pressure				
body weight	weight at birth	normal	high	Σ	% normal	Yules Q
≤ 90 kg	low	329	99	428	77%	-0.09
	high	221	55	276	80%	
> 90 kg	low	25	33	58	43%	-0.04
	high	107	131	238	45%	
Σ		682	318			

According to Tab. 2-16 there exist a slightly negative association between weight at birth and the actual blood pressure: Persons with a lower weight at birth tend to have a slightly increased blood pressure (the association is however quite low).
This relationship holds for people with a body weight ≤ 90 as well as for those with weight > 90.
The data in the marginal table (summed over the variable body weight) reveals quite a different picture:

Tab. 2-17: *Marginal contingency table resulting from Tab. 2-16 by summing over the values of the variable body weight.*

	blood pressure				
weight at birth	normal	high	Σ	% normal	Yules Q
low	354	132	486	73%	0.207
high	328	186	514	64%	
Σ	682	318			

For the data in the marginal table a positive relationship between blood pressure and weight at birth is observed: Persons with a higher weight at birth tend to have an increased blood pressure than those with a lower weight at birth.
Explain the observed discrepancy between the data in Tab. 2-16 und Tab. 2-17.
Comment:
The explanation should not consist in a description of the discrepancy but in a detailed explication of the reasons for the observed discrepancy.

Exercise 2-9:

Compute Yule's Q for the relationship between the variable *color of the victim* and *death sentence* for the data in Tab. 2-14:

(a) For the data in the full table conditional on the color of the delinquent.
(b) For the marginal table of the variables *color of the victim* and *death sentence* (that results from summing over *color of the delinquent*).

Note: The conditional values should be zero. However, due to rounding errors and the small frequencies in the table they are slightly different from zero.

Exercise 2-10:

Given:
The data of Ex. 2-32 (cf. the illustration in Figure 2-14).

Conduct the following analyses:

(a) The regression equation resulting from ignoring the hierarchical structure of the data.
(b) The regression equation that takes the clustered structure into account.
Comment: It is assumed that the same regression equation holds for each group.

(c) A test comparing the model assuming a single regression equation for the three groups with the model presuming three different regression equations for the three groups.

Present the following results:

(1) The result of the significance test concerning the regression parameters of (a).

(2) The result of the significance test concerning the regression parameters of (b).

(3) The result concerning the test of whether the assumption of a single regression line for each of the three groups is justified or not.

Exercise 2-11: *Lord's Paradox (Lord, 1967):*

Given:

The data of two groups (cf. Figure 2-19)

Illustrate Lord's paradox by:

(a) Conducting an analysis of covariance on the data of the posttest with the pretest as a covariate and with *school* as the independent variable.
This analysis should result in a significant effect of the factor school.

(b) Conduct a *t*-test for independent samples (according to the two types of schools) on the difference between pre- and posttest scores.
The results of the *t*-test should not reveal any significant differences both groups.

Comment:

Instead of an analysis of covariance a regression analysis may be conducted.

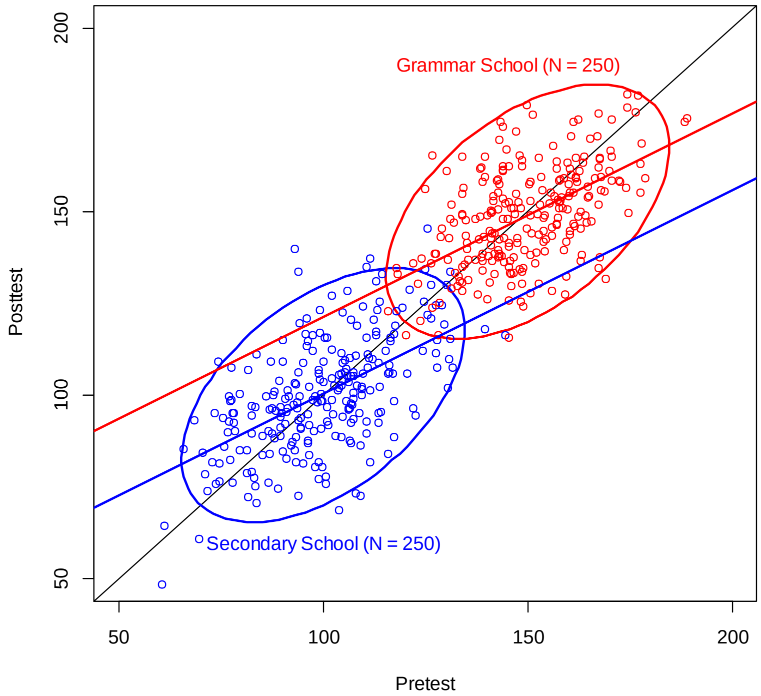


Figure 2-19: *Test scores of pupils from a secondary and a Grammar school (one year after shifting from the secondary school) [Simulated data]. The ellipses represent 95% confidence regions of bivariate normal distribution of the two populations from which the samples were drawn. The small circles represent the single data points. Also shown are the two (population) regression lines.*

2.9 Appendix: Paik Diagrams Illustrating Simpson's Paradox

This appendix presents diagrams of Paik (1985) to illustrate graphically Simpson's paradox for the examples presented in Section 2.6.3.
The *x*-axis represents the relevant independent variable, and the *y*-axis represents the relevant proportions (for the dependent variable).
The variable to be summed over is denoted by verbal labels. The proportion of observations within the different groups is indicated by the size of the (blue) circles.
The blue points and lines refer to the data of the full table, whereas the red points and lines indicate the proportions in the pooled sample.

Ex. 2-28: Simpson's paradox: Death sentences in Florida *during the years 1972-1979:*

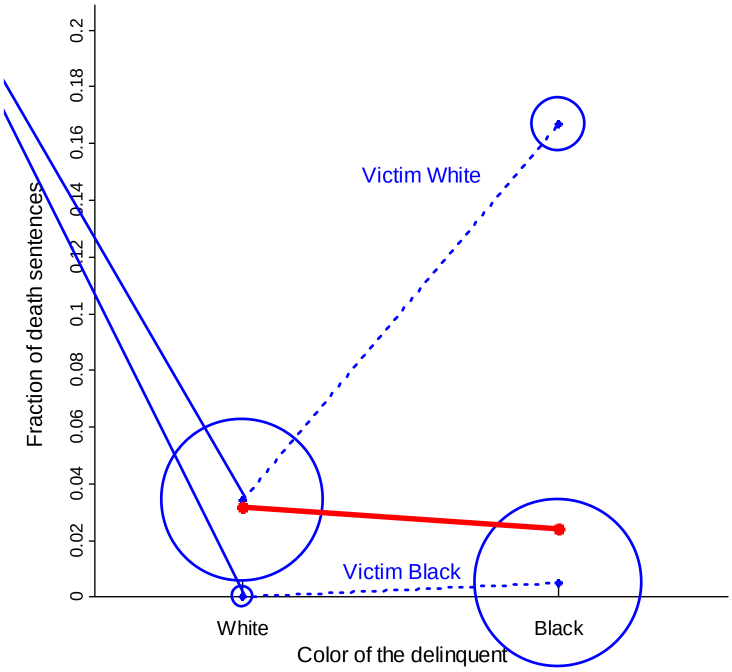


Figure 2-20: *Paik diagram illustrating the effect of summing over the relevant variable color of the victim for Ex. 2-28: Death sentences in Florida: The positive association between the death sentences and color of the delinquent within the full table (blue symbols) turns into a negative association when summed over the color of the victim (red symbols).*

Ex. 2-29: Simpson's Paradox: Treatment effects:

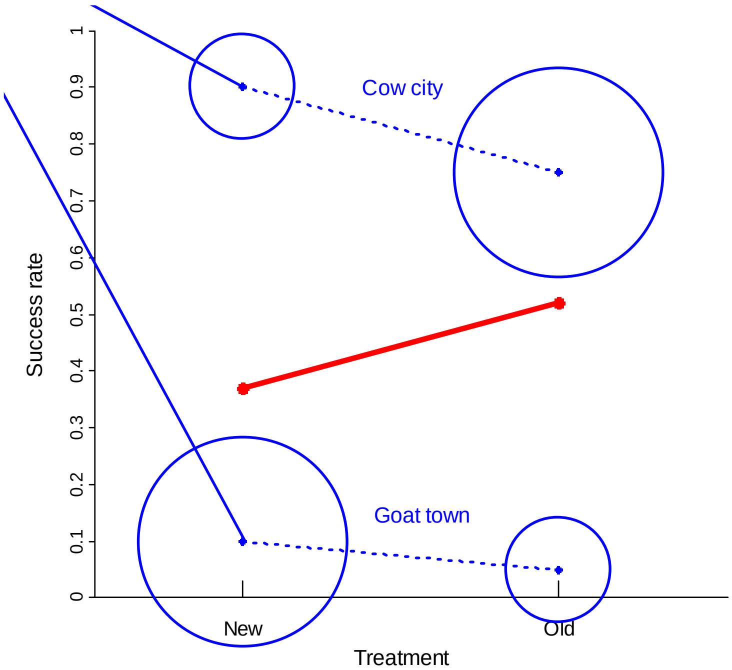


Figure 2-21: *Paik diagram illustrating the effect of summing over the relevant variable locality for Ex. 2-29: Treatment effects: The positive association between the new treatment and the success within the full table (blue symbols) turns into a negative association when summed over the locality (red symbols).*

Ex. 2-30: Simpson's paradox: Gender discrimination*:*

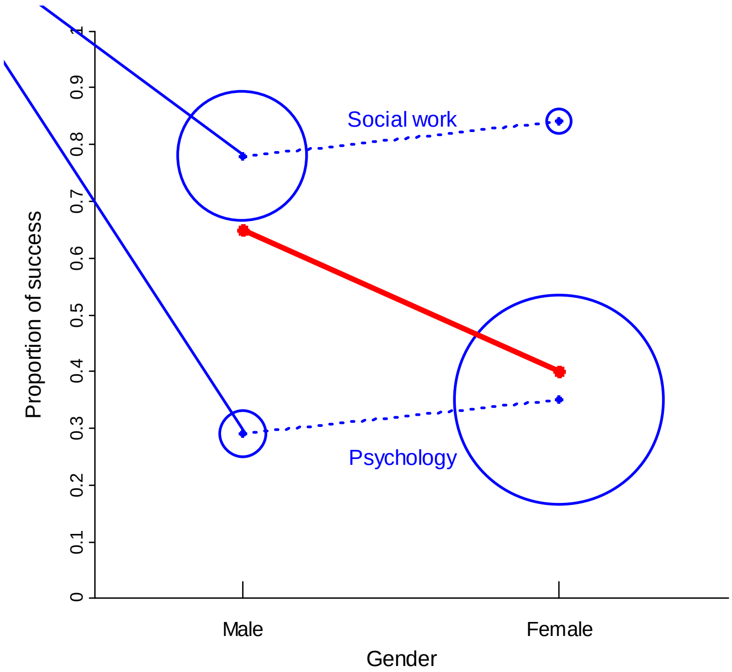


Figure 2-22: *Paik diagram illustrating the effect of summing over the relevant variable branch of study for Ex. 2-30: Sexual discrimination: The positive association between gender and study success within the full table (blue symbols) turns into a negative association when summed over the different branches of study.*

3. Memory Judgments

The present chapter considers various types of errors concerning memory judgments. The presentation starts with overview over classical errors of memory. This is followed by discussion of erroneous judgment concerning variations and constancy in personal life. Next we discuss the so called hindsight bias, a common type of error in everyday life. Finally a somewhat paradoxical phenomenon is discussed that concerns the retrospective estimation of painful experiences.

3.1 Classical Studies on Faulty Memories

The first known studies illustrating persons' inclination to adjust their memories in order to render a series of events more plausible come from Bartlett (1932). These studies indicate an important mechanism that has an impact on our memories.

Cognitive Mechanism 3-1: Encoding, retrieval, and forgetting of memory information

1. *Encoding of information:*

   The process of encoding of information should not be regarded a process of simply storing information. Rather the encoded information *is adjusted to the existing knowledge structure*. The process of encoding is thus influenced by the meaning of the information as well as by the existing knowledge structures.

   Usually the information is not encoded verbatim, i.e. in all its details. Rather the gist of the information is stored in memory. The following factors have an impact on the process of encoding:

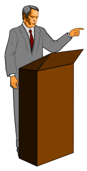
   - ❑ Attentional factors;
   - ❑ Depth of processing;
   - ❑ Existing knowledge;
   - ❑ Inferences draw;
   - ❑ Concomitant processing of related information;
   - ❑ Value and emotional meaning.

   Encoding should be conceived of as the construction and modification of existing of knowledge structures, respectively.

2. *Retrieval of information:*

   The retrieval of information from memory has to be considered as a *process of (re-) construction*: information of from different sources is integrated. These additional sources come from subjective theories and judgmental processes that are used to fill in gaps and to render the remembered content plausible and coherent.

3. *Forgetting / Interference:*
   According to the interference theory of forgetting the retrieval in memory may be blocked by interfering information. This can result in *intrusions*, that is, interfering information is retrieved thus blocking the retrieval of the target information.



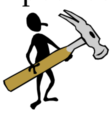
Comment 3-1:
The results of Bartlett could not be replicated. However, the cognitive mechanism presented is nevertheless generally accepted.

In the following this crude specification will be further refined by presenting more concrete mechanisms. To this end we focus on two lines of research that are not independent from each other. Both exerted a great impact on present memory research.

1. The studies of Elisabeth Loftus and colleagues concerning the influence of misleading information on the testimony of witnesses. An overview of this line of research may be found in Pickrell, Bernstein und Loftus, (2004).
2. The studies of Deese (1959) as well as of Roediger and McDermott (1995) concerning associative memory illusions. On overview of this line of research is found in Roediger and Gallo (2004).

3.1.1 Misleading Information and Memory Errors

In the seventies Elisabeth Loftus developed the method of misleading post event information.



Method 3-1: *Method of misleading post event information:*

The method of misleading post event information consists in an experimental procedure where systematic misleading information is presented to participants after the relevant event. The post event information was assumed to modify the memory of the original event.

Comment:
The misleading information consisted in certain questions that suggested the presence of certain events. The misleading information could not always be classified as being incorrect (Although same questions contained misinformation).

Example:
People saw a film about a traffic accident. Afterwards they were asked the following question (with participants in different groups received different wordings):

»About how fast were the cars going when they (smashed / collided / bumped / hit / contacted) each other?«

Loftus and colleagues demonstrated in a series of experiments that the presentation of misleading post event information can provoke wrong testimonies (e.g. Loftus & Palmer, 1974).
The following examples of everyday life demonstrate the effect of misinformation on memories. However, for these examples there was (with high probability) no corresponding real event. The information leading to the erroneous memories was thus not only misleading but simply false.

Ex. 3-1: Memory and suggestion (Loftus, 1980, p. 119):

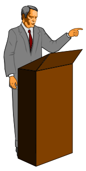
Jean Piaget reported that he had vivid memories of an attempt to kidnap him out of the baby carriage on the Champs-Élysée. He »remembered« the gathering of people, the scratches in the face of his heroic nurse that had saved him, the white push stick of the policeman, and the fleeing offender.
Despite their vividness Piaget's memories were wrong. Many years later the nurse confessed to have invented the whole story.

Ex. 3-2: Recollections of a deputy (Schacter, 1999, p. 214):

The 43 years old deputy Paul Ingram was accused by his daughters to have mistreated them sexually in their childhood.
Initially he denied vehemently all the accusations since he had absolutely no memory about these events. However, his colleagues, the officers, and the priest (Ingram was a member of a fundamentalist church) assured him that he will remember the events after having made a confession.
Following to long talks and examinations Ingram finally confessed and stated that he had probably repressed his memories of the events.
However, the officers believed that the sexual abuse has taken place in the context of Satanism, and, in fact, in the course of the examination Ingram »remembered« increasingly better the various events and accused further people that denied vehemently to have committed the criminal acts.
At the same time Ingram's daughters »remembered« further details of the Satanism like killing of babies, and a mass orgy.
On waiting on his lawsuit Ingram was visited and interviewed by social psychologist Richard Ofshe. The latter asked Ingram to remember how he has forced his son to have sexual intercourse with his daughter before his eyes.

Comment:
This event follows a logic that is similar to the other accusations. However, none of Ingram's daughters has ever claimed that such an event had occurred, and also the son denied it vehemently.

Ingram's reaction to the questions of the psychologist followed a predictable pattern: First he could not remember. However, following to intensive visualization and praying he developed vivid »memories« concerning the respective event.
Finally, Ingram was convicted of imprisonment for 20 years. At the time of Schacter's report (6 years after the conviction) he had not yet been released.



Comment 3-2:

1. The case of Ingram demonstrates that a confession does not implicate the guiltiness of the accused. For instance, more than 200 people confessed to have participated the kidnapping of the Lindbergh baby (On march, 1st, 1932 [Lilienfeld et al., 2010]).
2. Children are unqualified for serving as witnesses due to their inclination to confabulations (see, for example, Schacter, 1999).

The effect of misinformation and misleading information, respectively, has been widely demonstrated. However, there was disagreement concerning the explanation of the phenomena. Initially, Loftus assumed that the misleading information resulted in a replacement of the original memory by a new one.
Experiments of McCloskey & Zaragoza (1985) casted doubt on this interpretation for the following reason: If, in a final test, participants had to choose between the original and a modified picture the correct picture is selected about equally often independently of whether a misleading information was presented or not. If the misleading had resulted in a modification of the original memory trace then one would expect a higher rate of selection by those participants that were exposed to the misleading information.
A better explanation results from studies concerning *reality* and *source monitoring*.

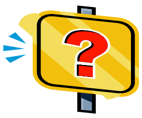
Cognitive Mechanism 3-2: Reality monitoring (Johnson & Raye, 1981):

Reality monitoring is concerned with the issue of how real events and imaginary events (dreams, thoughts etc.) are discriminated.

With respect to memories relevant cues for distinguishing correct from false memories were investigated: Vividness, imagination, level of detail, familiarity, plausibility, etc.

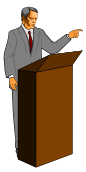
Cognitive Mechanism 3-3: Source memory and source monitoring (Johnson, Hashtroudi & Lindsay, 1993; Mitchell & Johnson, 2000, 2009):

The concept of *source memory* relates to memory phenomena that concern the source of information. Specifically, it is concerned with issues like: »Why is this person familiar to me?« or »Where did I get this information?«.
Source memory is much stronger affected by processes of aging than our memory of facts. In order to compensate for this deficit people refer to the strategy of plausible reasoning (Mather, Johnson, & DeLeonardis, 1999).



Question:
How can the findings concerning the effects of misinformation and misleading information be explained by means of source memory and reality monitoring?

Numerous studies demonstrate that the instruction to imagine in great detail specific events can lead to the impression that the imagined events have taken place in reality (e.g. Gerry, Manning, Loftus, and Sherman, 1996). This might explain why, following to »intensive visualization«, Paul Ingram revealed vivid memories of the events (cf. Ex. 3-2). In addition the assertions of Ingram and his daughters concerning satanic rituals were mutually confirming thus enhancing their plausibility and subjective probability, respectively. Both factors (visualization and plausibility) are able to subvert the capability to discriminate real from imagined events.
Source memory is important too since Ingram as well as Piaget had good memories with respect to the specific events. However, they had no memories with respect to the sources of the information. For example, Piaget was unable to remember that his vivid memories were due to the indoctrination of his nurse. He simply »accepted« the story of the nurse as part of his own experience.



Comment 3-3: Memory and therapeutic interventions

Take care of memories that are the result of various therapeutic techniques used to »recover« lost memories. In the United States these types of recovered memories have led to numerous condemnations because of sexual misuse.
Initially, Sigmund Freud believed that neurotic symptoms were the result of sexual misuse in childhood. Later on he abandoned this theory attributing reports of sexual abuse to patients' imagination.

In the eighties the discussion concerning the sexual abuse of Freud’s clients was heated up once again by Jeffrey Masson who indicated that Freud had swept the issue of sexual abuse under the carpet. However, Masson’s accusations could never be confirmed.

3.1.2 Associative Memory Illusions: The DRM-Paradigm

Another, more subtle way to provide misleading information was realized in the context of the co called *DRM paradigm* (Deese-Roediger-McDermott paradigm).

Method 3-2: *DRM Paradigm (Deese, 1959; Roediger & McDermott, 1995, Experiment 2)*

The *DRM paradigm* is an experimental procedure to elicit associative memory illusions.

Here is an example how the paradigm was realized in the second experiment of Roediger and McDermott (1995).

(i) *Experimental material:*

The experimental material consisted in 24 lists of 15 words each. For each list there existed a critical word that was closely associated with the words on the list. However, critical words themselves did appear on the lists.

Example:

Critical word: *sleep*

The corresponding word list comprised the following 15 items:

Bed, rest, awake, tired, dream, wake, snooze, blanket, doze, slumber, snore, nap, peace, yawn, drowsy.

(ii) *Procedure:*

16 of the 24 lists were presented, one list after another with a rat of 1.5 seconds per word.

For half of the lists a free recall followed to the presentation of a list. Participants were instructed to recall as many words from the list as possible *without guessing*.

For the other 8 lists a short distractor task (solving of math problems) had to be performed after the presentation of a list.

Following to the presentation of the 16 lists (with the assigned tasks) participants had to perform a recognition test: 96 words from the following categories were presented:

❑ 48 new words (i.e. word not on the lists) including the 16 critical words.

- ❑ 48 old words with 24 words from the 8 lists followed by a recall task and 24 words from the list with a subsequent math task.

For each word presented in the recognition test participants had to indicate:

a) whether the word was old or new, and
b) in case of an »old« response, whether they could remember a concrete detail of the episode in which the word was presented (=Remember response) or not (=Know response).

(iii) Principal results:

- ❑ Critical words were reproduced during the free recall about equally often as words on the lists.
- ❑ In the recognition test slightly more critical words than words actually on the list were categorized as old.
- ❑ Of the critical items actually produced in the recall task a remember response was given in 73 percent of the cases, i.e. participants believed that they recalled some details of the presentation episode.
- ❑ If the critical item was not produced in the recall task the rate of remember response was only 38 percent.

The following cognitive mechanism is required for explaining the phenomenon.

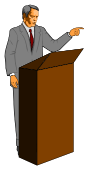
Cognitive Mechanism 3-4: Implicit associative response

The presentation of an associated item leads to the activation of the critical word.

The explanation of an implicit activation of associated items is evidenced by the following results:

- ❑ The false memory effect is also observed with rapid presentations (40 ms) of items.
- ❑ In a follow-up task priming effects on implicit memory measures were observed.

The implicit activation of the critical word increases its *fluidity*. This increases the familiarity of the word. The increase of familiarity might not be consciously realized by the person.
On the other hand the implicit activation may increase participants' tendency to recall the word consciously.
Both factors, the increase of familiarity as well as conscious recollection of the critical word during learning, increases participants' tendency to reproduce the word in a later recall test.



Comment 3-4: Priming effects and tests on implicit memory

In general *priming effects* are effects of facilitation: Due to processing of information the processing of subsequent information that is associated to the previous one is facilitated.

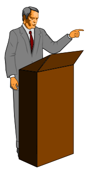
For example, in *semantic priming* the lexical decision concerning a word (whether a presented item is a word or not) is facilitated (as evidenced by faster reaction times) by a prior presentation of an associated word.

Tests on implicit memory are not recognizable as memory tests. They are used to measure effects on memory that cannot be readily assessed by means of explicit measures (i.e. recall and recognition).

A word completion task constitutes an example of an implicit test: It is tested whether a word is completed to a previously presented word or a word that is associated with a previously presented one.

The process of implicit associative responses can explain the high rate of »recall« of critical items. However it does not explain the high rate of remember responses for critical words. To explain this phenomenon *reality* (cf. Cognitive Mechanism 3-2) as well *source monitoring* (cf. Cognitive Mechanism 3-3) have to be taken into account.

The process of reality monitoring is relevant in the actual context since repeated implicit activations of the critical items enhance the fluidity and familiarity of the item. Both aspects constitute cues for assessing the degree of reality of a memory. By consequence a high level of familiarity and fluidity of items can lead to an illusion of reality.



Comment 3-5:

The significance of monitoring processes is supported by neurophysiological data: People with lesions in the frontal part of the cortex reveal a higher DRM effect. As is well-known, the frontal cortex plays an important part in control and executive processes.

This aspect also explains the finding that memory illusions can be elicited relatively easy in children since in this population the control and executive processes are not yet fully developed.

Perhaps, a still more important role is played by the process of source monitoring. Remember that remember responses to critical words were produced predominately for those critical words that were erroneously recalled in the recall test. Seemingly, participants remember to have seen the word previously. However, the information about the source (i.e. whether the item was presented on the list or whether it has been written down in the recall test) has been lost.

3.2 Judgments Concerning Stability und Change

The results presented in the previous section demonstrate the significance of control and monitoring processes for memory judgments. Moreover, Ex. 3-2 indicates the effect of processes of inference on memory judgments.

The effect of plausible inferences on memory has been nicely illustrated in studies concerning the judgments about stability and change in personal life.

Ex. 3-3: Biases in judgments concerning past attitudes (Marcus, 1986):

Objective of the study:

The objective of the study consisted in an examination of political attitudes and their changes between the years 1965-1982.

Participants:

1669 participants (Students and their parents); 64 percent participated the questioning at each of the three time points.

Procdure:

Participants were interrogated about their political attitudes at three points of time: 1965, 1973, and 1982. Specifically, they had to specify on 7-point scale their attitudes to each of the following 5 issues:

- Job guarantee,
- Rights of accused persons,
- Support of minority groups,
- Legalization of marihuana,
- Gender equality.

In addition they had to classify their political attitudes as being either conservative or liberal.

Results:

The estimation in 1982 of their own attitude from the year 1973 revealed nearly identical values to their present attitude (of the year 1982). A linear regression with the actual attitude (= *E*82) and the attitude of the year 1973 (= *E*73) as independent variables and the retrospectively assessed attitude of the year 1973 (*RE*) as dependent variable:

$$RE = a + b_{E82} \cdot E82 + b_{E73} \cdot E73\,,$$

Revealed a substantial higher weight for *E*82 than for *E*73. Ostensibly, the actual attitude had a higher impact on the retrospectively judged attitudes of 1973 than those attitudes (in 1973) themselves.

A process of anchoring and adjustment can be used to explain the results.

Cognitive Mechanism 3-5: Anchoring and adjustment

In general, the heuristic of anchoring and adjustment is used for making quantitative estimates.

The heuristic of *anchoring and adjustment* consists in two steps:

1. In the first step an initial estimate is generated on the basis of available information. The available information constitutes the anchor.
2. In the second step the initial estimate is adjusted due to further considerations based on subjective theories.

The main problem of the method of anchoring and adjustment consists in the fact that quite irrelevant information can constitute the basis for the initial estimate. Moreover, the process of adjustment is not sufficiently achieved. The interplay of these two factors can result in biased judgments.

Ex. 3-4: Anchoring and adjustment (Tversky & Kahneman, 1974):

In the experiment a wheel of fortune was spun first. Subsequently the following question presented to participants:

How high is the percentage of African nations in the United Nations?

For the group with the needle of the wheel of fortune indicating the number 65 the median estimate was 45 percent.
For the group with the needle of the wheel of fortune indicating the number 10 the median estimate was 25 percent.
Comment:
The median denotes the value with 50 percent of all values being located below and 50 percent being located above.

Interpretation:
Participants used the value indicated by the wheel as an anchor, and then adjusted (insufficiently) their estimate (up and downward, respectively).

The behavior of participants can be explained by means of anchoring and adjustment in the following way: In order to assess their attitudes in 1973 their actual attitudes were used as an anchor. Subsequently an adjustment of initial estimates was performed. However, the results of the regression analysis demonstrate that the process of adjustment was performed in an insufficient way only.
A detailed analysis of the estimated attitudes reveals that the process of adjustment was guided by means of plausible reasoning. For exam-

ple, people judged (erroneously) their attitudes as being more liberal in 1973 than in 1982, or parents estimated their attitudes as being more stable than those of their children. This corresponds to the common opinions that »with increasing age people become more conservative« and »elderly persons are more stable in their attitudes«.
The results of Marcus (1986) are confirmed by further studies revealing that the observed biases also concern other personal attitudes (e.g. Ross, 1989). Of further interest is the study of Conway & Ross (1984) illustrating how results confirming expectations can be achieved by means of biasing the own memories.

Ex. 3-5: Biased memories concerning the own performance (Conway & Ross, 1984):

Objective of the study:
The goal of the study consisted in assessing the effect of a University program to increase students' learning capabilities.

Procedure:
Participants were randomly assigned to one of two groups: Participants of experimental group took part in the program whereas participants of the control group were put on a waiting list.
The effect of the program was evaluated by means of a pre and posttest.

Results:
- The program had no effect.
- Participants of the experimental group estimated their performance prior to the training as inferior compared to students of the control group.

Comment:
A similar result was observed with respect to treatments of pain. Clients subjected to a therapy tend to overestimate the severity of their pain prior to the treatment (Linton & Mehlin, 1982).

The results presented demonstrate the influence of subjective theories on judgments of stability and change with respect to the own personal history.

3.3 Hindsight Bias

Let us start with a specification of the concept.

Concept 3-1: *Hindsight Bias*

Hindsight bias consists in an erroneous assessment of their prior knowledge concerning specific events. Specifically, in hindsight, people usually suppose a higher agreement between their knowledge prior to the relevant event and their actual knowledge (after the event) than is actually the case [»I knew it all along«].

Hindsight bias is quite a common phenomenon. Here is a classical study:

Ex. 3-6: Hindsight bias (Fischoff & Beyth, 1975):

The study comprised two phases:
In the first phase various groups of students were ask to provide their subjective probabilities for 15 events concerning Richard Nixon's visit of the USSR and China in the year 1972.

Examples:

- ❑ Will the United States install a diplomatic department in China?
- ❑ Will Nixon initialize a common outer space program with the USSR?

In the second phase (2-6 weeks after Nixon's visit), participants had to remember their estimates provided in Phase 1. In addition, they should indicate whether they believed that an event had actually occurred or not.

Results:

1. 3/4 of the students »remembered« higher estimates than provided in the first phase for those events they believed to have actually taken place.
2. Most students »remembered« lower estimates than provided in the first phase for those events they believed not to have occurred.
3. Hindsight bias increases with the distance to the actual events: After 3-6 month 84% of the participants exhibited hindsight bias.

Hindsight bias might also be explained by means of *anchoring and adjustment*: The actual outcome functions as an anchor and persons are unable to adjust their judgment readily in order to correctly assess their predictions given previously.
Hindsight bias can be reduced by having people search for possible reasons why an occasion might have resulted in a different outcome. In this case, people are better apt to distance themselves from the anchor and to provide a more appropriate adjustment.

3.4 Retrospective Evaluation of Affective Episodes

A series of studies demonstrates that, in their retrospective evaluation of past positive or negative episodes, people do not put sufficient emphasis on the duration of an episode. This is nicely illustrated be the subsequent experiment.

Ex. 3-7: Restrospective evaluation of painful experiences (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993)

Participants were subjected to two different (and slightly painful) experiences with cold water under pressure, one for each hand.
In the *short episode* the hand was dipped into cold water of 14°C, for 60 seconds.
In the *long episode* the hand was first dipped into cold water of 14°C, for 60 seconds. This was followed by a period of 30 seconds during which the temperature of water was gradually increased from 14°C to 15°C. The hand remained in the cold water during the whole episode.
After a short time interval participants were asked to indicate which of the two conditions they would prefer to repeat once again.

Result:
Paradoxically significantly more participants preferred to repeat the long episode despite the fact that this episode includes the short one.

The result of this study was applied in a medical context.

Ex. 3-8: Retrospective evaluation of the experience of a colonoscopy (Redelmeier, Katz, & Kahneman, 2003)

Participants (N = 682) consisted of clients that were subjected to a colonoscopy.
For half of the participants the intestinal tube was not removed immediately following to the medical examination. Rather it remained for about 1 further minute in the bowel without being moved (=prolonged examination).

Results:

- ❑ Participants of the group with a prolonged examination provided, in general, a more positive retrospective evaluation of the examination than participants of other group.
- ❑ In an interview 5 years later participants of the group with prolonged examination showed a higher willingness to repeat the procedure compared to participants of the other group.

In order to explain the observed phenomenon the following mechanism was proposed.

Cognitive Mechanism 3-6: The »snap shot model« of retrospective evaluations of affective episodes (Fredrickson & Kahneman, 1993)

Basic model assumptions:

1. The episode to be evaluated retrospectively is represented in memory by prototypical moments.
   In case of affective episodes the representation consists of moments of increased affective value.
2. In addition, the final phase of the episode is registered thus making up a part of the representation of the episode too.
3. The retrospective evaluation consists in a combination (a weighted mean) of the evaluation of the moments of the affective summits and the evaluation of the affect at the end of the episode.

In case of Ex. 3-7 and Ex. 3-8 the prolonged episodes are more positively evaluated because of the less negative segments at the end of the whole episodes.
It is important that the less negative or more enjoyable phase at the end of the episode must be perceived as a part of the episode. A positive or rewarding event following to the episode need not result in a higher evaluation of the episode itself since it may not be regarded as part of the episode.

3.5 Chapter Summary

In the present chapter the following phenomena and cognitive mechanisms were discussed:

- ❑ The (repeated) presentation of misinformation or of misleading information can result in false memories that are experienced as vivid and realistic.
- ❑ Suggestive and imaginative techniques to »regain« lost memories, as used in various treatments, are particularly apt to evoke vivid false memories. This is due to the negative influence of these treatments on processes of reality and source monitoring: During the therapy cues are generated repeatedly pretending erroneously reality. In addition, with repeated application judgments concerning the source of the information leading to false memories are made more difficult.
- ❑ A specific way to present misleading information consists in the presentation of information that is either semantically related to a target or closely associated to it (as in the DRM paradigm). The associated or semantically related information results in implicit activation of the target items leading to associative memory illusions.

- In the process of remembering memory gaps are filled be means of plausible inferences. This guarantees the internal consistency as well as the plausibility of the remembered stuff. However, it also can result in faulty reinterpretations of past state of affairs.
- Hindsight bias consists in a biased evaluation of own predictions about subsequent events in order to achieve a higher consistency between predictions and reality.
- Anchoring and adjustment is a cognitive mechanism that might be used for assessing previous judgments and predictions: The actual knowledge is used as an anchor for the assessment process. An additional adjustment process based on plausible inferences is performed in order to arrive at a more realistic judgment. In many cases the process of adjustment is insufficient resulting in a biased assessment.
- The retrospective evaluation of negative episodes incorporates information about the length of an episode in an insufficient way. This can results in the paradoxical phenomenon that longer negative episodes that contain a shorter one as its part are evaluated as more positive than the shorter one.
- The »snap shot model« of retrospective evaluation provides an explanation of this phenomenon by assuming that only specific moments of the whole episode as well as the final segment are registered.

4. Probability Judgments and Probabilistic Inferences

The present chapter is concerned with probability judgments. Biases of probability judgments can be assigned to four broad classes:

(a) The probabilistic nature of a phenomenon is ignored or not adequately taken into account (For example, the phenomenon of base rate neglect discussed in Chapter 4.3.2).

(b) The probabilities assigned to single events reflect the actual proportions only partially. This sort of bias indicates that the probability judgment is influenced by sources other than the actual frequency of occurrence (cf. the effects of the avaliability and representativeness heuristic discussed in Chapter 4.2).

(c) Inconsistent assignment of probabilities to different events: The assignment of probabilities to two or more events contradicts the axioms of probabilities (cf. the conjunction fallacy discussed in Chapter 4.2.2).

(d) Errors in probabilistic reasoning: The inference of probabilities of specific events on the basis of the given probabilities does not correspond to the laws of probability (cf. Chapter 4.4).

In each of these cases subjects are *unable to process proper representtations of the problem situation* resulting on the observed biases.

The issue of biases in probabilisitic judgments has received and still gets great attention. In addition there have been numerous debates concerning the validity and the interpretation of the empirical findings.

The chapter starts with a discussion of basic concepts of probability. This should result in a better understanding of the subsequent discussion of errors in probabilisitic judgments and reasoning.

The second section presents the *heuristics and biases* approach of Tversky and Kahneman that had and still has an enormous impact on the research in the field.

The third subchapter discusses various errors and biases concerning the use and interpretations of probabilities, like *base-rate neglect*, the erroneous treatment of conditional probabilities, and the misinterpretation of information about risks in medicine.

The fourth part is concerned with probabilistic inferences. We focus on investigations in the context of *Bayes theorem* since a great deal of investigations on probabilistic reasoning are more or less related to Bayesian reasoning. This section discusses the principles of probabilistic inference in general, and it will be shown that Bayes rule is just a specific application of these principles of probabilistic reasoning.

The fifth section discusses dual-process theories which have recently received some attention as possible expanations of judgmental biases.

The next section presents different methods on how to improve probabilistic reasoning.

Finally, we discuss various criticisms concerning the investigation on biases of probabilistic reasoning.
The presentation comprises three aspects:

(a) Empirical findings concerning biases and errors in probability judgments.
(b) Psychological theories and mechanisms for explaining the empirical findings.
(c) Normative and methodological issues concerning basic concepts of probability theory and methods of probabilisitic reasoning (The appendix *Elements of Probability Theory* provides a more in depth discussion of the normative issues).

4.1 Conceptions of Probability and the Meaning of Probability Statements

Probability theory can be conceived of as a basic sciences that is relevant to all types of theoretical and applied sciences, like physics, economics, medicine, and the social sciences (to name but a few). In addition, in daily life people are increasingly confronted with statements concerning the probability of different events, like death statistics, crime rates, etc. Due to this growing significance, elementary probability theory and statistics have found their way into the curricula of most high schools. Moreover, the issue of how to best teach probabilistic thinking and reasoning has become the subject of scientific research (cf. Chernoff & Sriraman, 2014). In the present section, different conceptions of probability and the meaning of probability statements are discussed.
Despite the fact that probability computations are important for most sciences there does not exist a unique uncontroversial conception of probability. In fact there exist different views resulting in different interpretations of probabilities. In the following the two most important conceptions, the *objective* and the *subjective* or *Bayesian* conception are presented.

4.1.1 Objective Conception of Probability

According to the mathematical conception of probability, given by the probability axioms of Kolmogorov, probabilities measure the size of sets. Probability is thus called a *normed measure* that has the same general characteristics as other measures like length, volume, weight, counts, etc. The probability measure is normed in that the maximal possible value is 1.0.
The objective conceptions of probability conform closely to the mathematical conception since probabilities are assumed to apply to *sets*, *ensembles*, *populations*, etc. of existing things (therefore the characterrization *objective*).

According to the *classical conception* of probability that may be conceived of as a special case of the objective conception the probability of a target event corresponds to the proportion of cases with this event being present within the target population. By consequence, probability statements refer to populations and not to single units.

Ex. 4-1: Interpretation of probability statements within the classical conception of probability

The following statement:

> *The chance of getting a cardiovascular disease is about 12% for a German man and about 6.5% for a German woman (Müller, 2019).*

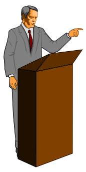
does not refer to a single man or woman but to the German population of men and women, respectively.

However, the statement may be interpreted as follows:

> *Drawing a man at random from the population of German men, the chances that he gets a cardiovascular disease are about 12%.*
> *Similarly, drawing a woman at random from the population of German women the chances that she gets a cardiovascular disease are about 6.5%.*

Thus the statement might also be interpreted as applying to a *randomly drawn person* from the respective population.

It is importan to realize that this does not mean that the probability may be assigned to a *concrete single person* since a concrete person either gets the disease or not (there is no probability involved). Thus the modifier *randomly drawn* is important and must not be dropped.



Comment 4-1: On the problem of drawing at random sample from a population

The latter interpretation of probability in Ex. 4-1 has a problematic aspect that might go unnoticed at first: How can we draw at random from a population?

In practice there exist many random devices, e.g. statistical packages implement different random generators that can be shown to exhibit a good random behavior. However, these devices are not truly random. By consequence they are often called *pseudo-random generators* that closely simulate random devices

> Most random devices, we know of, exhibit some sort of serial dependence which might be quite small such they may count as acceptable approximations to random devices. This also applies to such devices as throwing a (fair) coin or Casino games, like roulette.
> In fact, we do not know whether there exists something like a real random device.

In summary, according to the classical conceptions probability statements always refer to some sort of population. Specifically, it refers to the proportion of target cases within this population. In this way the classical account conceptualizes probability as a count measure that is characterized by the probability axioms.
The classical conception of probability is due to Pierre-Simon Laplace (1749-1827) and is based on the idea that each member of the population has the same chance of being selected. Thus it implicitely assumes equiprobability of the elements of the popualtion. A different approach within the objective conception was taken by Richard von Mises (1883-1953). The basis of his *frequentist conception* is the random experiment, like tossing a coin or throwing a die that may be repeated infinitely often. The different outcomes of the experiment need not be equiprobable. The probability of an event E is defined as the limit of N_E, i.e. the number of cases that E appears within the resulting sequence, divided by the length N of the sequence, i.e. the number of repetitions of the random experiment as N goes to infinity. Thus the probability is not a proportion of positive cases to the whole number of cases as in case of classical probability but *the limit of the proportion of positive case within the whole sequence as the latter approaches infinity*.
The classical conception of probability might be conceived of as a special case of the frequentist conception with the random experiment consisting in random sampling with replacement from a given population. With increasing sample size the proportions in the sample reflect more closely the true proportion in the population, i.e. the true probabilities. There is, however, an important difference: with finite populations the probability can in principle be observed by examining the population. Consequently, the probability is an *observable quantity*. By contrast, in case of flipping a coin, throwing a dice, etc. the probability cannot be observed directly. It is thus a *theoretical quantity* that can be measured with error by investigating the sequences of outcomes.
A third objective conception of probability is Karl Popper's (1902-1994) *propensity conception* of probability (Popper, 1959). According to this approach probability refers to a *probabilistic setup*. For example, in case of flipping a coin, the probabilistic setup consists in the coin and the flipping device. Note that the probabilistic setup does not

consists in the coin only but comprises all factors that are relevant for how the coin lands. Similar to the frequentist conception probability is assumed to be a theoretical quantity that can only be measured erroneously by running the setup.

To summarize, according to the objective conception probabilities are associated with exisiting things like populations, sequences (though infinite sequences do not really exist) or physical probabilistic setups that are able to produce (infinite) sequences. This contrasts with the subjective conception according to which probability is linked to our knowledge.

4.1.2 The Subjective Conception of Probability

According to the objective conception the following statements do not make sense:

- *The probability that the universe has its origin in the Big Bang is 0.90.*
- *The probability that the extinction of the dinosaurs was cause by a meteor strike is 0.60.*
- *The probability that the YB Bern will win the next Swiss soccer championship is 0.50.*

In each of these cases the probability does not refer to a characteristic of an objective physical event. Instead, the probability refers to subjective knowledge and background assumptions, respectively. Note that the events in question either have occurred (or not) or they will occur (or not). Thus there is no objective probability of occurrence. The uncertainty expressed by using the term probability is due to our lack of knowledge: In case of perfect knowledge there would be no room for probabilities. Thus, according to the *subjective conception* of probability, the term refers to our subjective knowledge and assumptions and not to real events.

The fact that the probability depends on the subjective knowledge and assumptions implies that two different people may assume different probabilities concerning the occurrence of a specific event. However, the subjective probability may be revised by incorporating objective evidence about the occruence of an event resulting in a convergence of their subjective probabilities. A simple example illustrates this point.

Ex. 4-2: Convergence of subjective probabilities

Given:

1. Two hypothesis concerning the probability of a coin landing Head or Tail:

H_1: The coin is fair: $P(\text{Head}) = P(\text{Tail}) = \frac{1}{2}$.

H_2: The coin is biased: $P(\text{Head}) = \frac{1}{3}$, $P(\text{Tail}) = \frac{2}{3}$.

2. Two persons P_1 and P_2 have the following assumptions about the probability of the two hypotheses:

   P_1: $P(H_1)=\frac{9}{10}, P(H_2)=\frac{1}{10}$.

   P_2: $P(H_1)=\frac{1}{2}, P(H_2)=\frac{1}{2}$.

Thus, our two subjects have different subjective probabilities concerning the fairness and biasedness, respectively, of the coin.

Now the coin is flipped 100 times and the number of heads that show up is 48.

Assuming that the single trials are independent and using the optimal updating rule (Bayes rule), the probabilities of the two hypotheses for the two subjects after taking the evidence from the coin tossing experiment into account are:

   P_1: $P(H_1)=0.998, P(H_2)=0.002$.

   P_2: $P(H_1)=0.978, P(H_2)=0.022$.

The subjective probabilities of the two persons have converged considerably. With increasing number of trials (more data) the subjective probabilities become closer and closer thus the initial subjective probabilities become more and more irrelevant.

Comments:

- ❑ Bayes rule will be discussed in great detail in Chapter 4.4.
- ❑ An additional example illustrating the updating of subjective probabilities is provided in Ex. 4-15, on page 130.

Ex. 4-3: Classical vs. Bayesian statistics

In statistics there exists a divergence in opinion concerning the nature of the parameters characterizing a population.

According to the *classical (frequentist) view* population parameters (like means, proportions) that are used to characterize populations have to be conceived of as fixed quantities (constants). The main objective of statistics consists in estimating the values of the parameters, on the basis of the given data, and to draw inferences about the values of the parameters.

In *Bayesian statistics* population parameters are conceived of as random variables that follow a specific distribution, called the *prior distribution*. The data are used to revise the prior distribution and to arrive at a new distribution that takes the empirical evidence into account. The resulting distributeon is called the *posterior distribution*. On the basis of the posterior distribution inferences about the underlying population can be drawn.

Comments:

1. Bayesian statisticians do not necessarily reject the idea that population parameter (e.g. the mean size of the women in the Swiss population at a give time point) are fixed quantities. However, the uncertainty due to our lack of knowledge should be taken into account. By consequence, our knowledge of the parameter is better represented by a distribution that covers the plausible values, e.g. a normal distribution with mean, say, 1.70 and a standard deviation of 0.15 as the mean hight of Swiss women at age 20.
2. With large samples the difference between classical and Bayesian statistics vanishes. This is due to fact that the the data dominate the prior assumptions. Consequently, the posterior distribution is determined predominantly by the data and not by the prior assumptions (cf. Ex. 4-2).

Let us summarize our considerations of the different types of conceptions of probability:

1. In order to cover different sorts of probabilistic statements one requires at least two conceptions of probabilities: objective and subjective probabilities.
2. According to the objective conceptions probabilities refer to sets of exisiting things or to generating mechanisms of these sets. The subjective conception links probabilities to subjective knowledge.
3. Within the objective conception, the true probabilities are *estimateed*. Existing estimates may be improved by including further data (e.g. further trials of a random experiment or a greater sample from the population).
4. Within the subjective conception, subjective probabilities may be revised due to further information. An important source of information (but not the only one) consists in the inclusion of empirical data.

4.2 The Heuristics and Biases Approach of Tversky and Kahneman

The heuristics and biases approach was initiated in the seventies by Amos Tversky (1937-1996) and Daniel Kahneman (1934-). These investigations were honored by granting the Nobel price of economics to Kahneman in 2002.

The basic idea underlying the heuristics and bias approach consists in the assumption that a great deal of our probability and frequency judgments, respectively, is based on heuristics. Heuristics are rules of thumb that may either result, with relatively little effort, in satisfying inferences or in biased judgments, depending on the specific circumstances.

The usage of heuristics may be conceived of as an answer to the problem of *bounded rationality*. This concept is due to Herbert Simon (1916-2001) and refers to humans' bounded information processing capacities that prevent optimal decisions and judgments even in case of complete information. Bounded rationality induces people to search for satisfying instead of optimal decisions. Simon calls this process *satisficing* in contrast to the process of *optimizing*. In addition, more recent research has revealed that maximizers tend to be less happy than satisficers (Pohlman, 2010). This is due to various reasons, for example, additional costs for finding an optimal solution are not worth the received gain. A further reason consists in the fact that expectations of maximizers are often too high thus resulting in frustration.

The results of the research of Herbert Simon as well as of Tversky and Kahneman have long been ignored by in the economist community that had »dreamed« of the perfect rational individual for a long time. However, the results have finally found some acceptance in this community leading to research of economic behavior that is based on the investigations of Tversky and Kahneman.

The investigations of Tversky and Kahneman were concerned primarily with judgmental errors resulting from the application of heuristics. However, they also present examples where heuristics provide good results. The three most important heuristics may be summarized as follows:

Cognitive Mechanism 4-1: Heuristics of the heuristics and biases approach

There are three focal heuristics whose effects on human judgments have been investigated in great detail:

1. *The availability heuristic:*
   The availability heuristic is based on the generation of instances of cases from the target class whose frequency or probability has to be assessed. The more cases come to mind (or the easier the task of generation of cases) the higher the probability (or frequency) estimate of the target class.
2. *The Representativeness heuristic:*

The representativeness heuristic is based on the typicality (or similarity) of an instance with respect to the target class: The higher the typicality of the instance the higher the estimated probability that the instance is a member of the target class.

3. *Anchoring and adjustment:*

The heuristic of anchoring and adjustment has been described above (cf. Cognitive Mechanism 3-5: Anchoring and adjustment, on page 98).
As demonstrated) above in Ex. 3-4 on page 98 the *availability* of information is also important in the application of the heuristic of anchoring and adjustment since available information has an influence on the generation of the anchor even if this information may be completely irrelevant with respect to the event whose frequency has to be assessed.

In the following the mode of functioning of the availability and representativeness heuristic will be illustrated by means of various examples.

4.2.1 The Functioning of the Availability Heuristic

Tversky and Kahneman (1973) illustrated the functioning of the availability heuristic in the context of frequency judgments in a series of experiments. Here are two examples:

Ex. 4-4: Functioning of the availability heuristic I (Tversky & Kahneman, 1973):

Consider the arrangements in Figure 4-1. Which arrangements allows for more different paths?
A path is defined as a line connecting a diamond in the first row to one in the last row where the line passes exactly through one diamond in each row (cf. Figure 4-1).

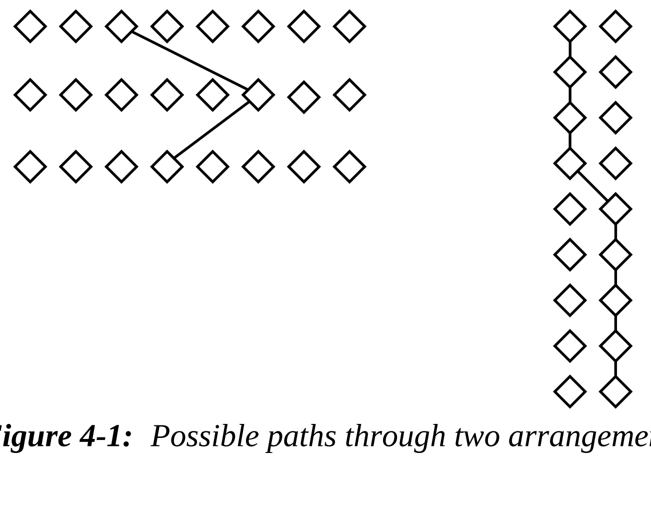


Figure 4-1: *Possible paths through two arrangements.*

Most people suspect that the left arrangement allows for more different paths. In fact, both structures permit the same number of paths, since $8^3 = 2^9$.

This example does not seem to be very realistic since it is quite irrelevant whether a person is able to assess the number of path in the two arrangements.

However the example illustrates how people try to solve the problem of estimating frequencies for a specific class of judgment problems if they miss a method for finding an exact solution: Due to the impossibility to count all possible paths within the given time limit people try out a number of different possibilities for both structures. It seems to be much easier to generate different paths for the left arrangement (due to a greater number of starting positions) resulting in a higher frequency estimate for this structure.

Let's look at a more realistic example:

Ex. 4-5: Functioning of the availability heuristic II (Tversky & Kahneman, 1973):

Participants:

The sample was balanced with respect to sex. Thus half of the participants were women and the other half were men.

Material:

2 lists, *A* and *B*:

List *A* contains famous names of 19 women and common names of 20 men.

List *B* contains famous names of 19 men and common names of 20 women.

Procedure:

The experiment was labeled as a memory experiment. The task of participants consists in memorizing as many names as possible.

Half of the participants received List *A* whereas List *B* was presented to the other half of the participants.

Names on the lists were presented in random order. They were read off the list in normal speed.

At the end of the experiment participants had to estimate the number of names of women and men on the lists.

Results:

- The sex comprising the famous names was judged as being more frequent independently of whether the members of the respective sex were men or women.

- ❑ In general, the frequency of the number of unknown names is underestimated whereas the frequency estimate of the number of famous names is close to the correct answer.

Interpretation:

The relevance of the availability heuristic for the results presented is obvious: The more famous names are better memorized and have thus a higher availability. By consequence they are judged as being more frequent.

The availability of information is heavily determined by the public media. Due to the selectivity of reporting the availability heuristic predicts biases with respect to frequency of various target events.

Ex. 4-6: Effect of the availability heuristic: Assessment of the rates of various causes of death

Please select from each pair of possible causes of death the one with the higher frequency (with respect to the United States of America):

- ❑ Being killed by falling aircraft parts vs. being killed by an attacking shark.
- ❑ Diabetes vs. murder.
- ❑ Car accident vs. stomach cancer.

Answers (From Death Odds, September 1990):

- ❑ The probability of being killed by falling aircraft parts is about 30 times that of being killed by an attacking shark.
- ❑ The number of people dying from Diabetes or stomach cancer is about twice the number of people being murdered or dying in a car accident.

The real odds of the different causes of death are surprising to many people. This can be explained easily by means of the availability heuristic: The »silent killers« like myocardial infarction or apoplectic stroke are not worth being reported in the public media whereas attacks of sharks or car accidents find their ways into the media.

4.2.1.1 Imagination and Availability

In Section 3.1.1, it was argued that specific imaginative techniques can result in false memories. It seems natural to assume that the intense imagination of specific event might lead to an increase of the subjective probability of the respective event. This hypothesis was tested in a study of Carroll (1978).

Ex. 4-7: Imagination and availability (Carroll, 1978):

The experiment took place on the day of the US presidential elections in 1976 (Jimmy Carter vs. Gerald Ford).

The experiment comprised two conditions depending on the content of the scenarios that have to be imagined:

1. Ford wins the elections as Carter fails to hold some key states and Ford wins in the Midwest and West. He wins 316 electoral votes whereas Carter wins only 222. Furthermore, Ford wins in 32 states whereas Carter gains the majority in only 22 states plus the District of Columbia.
2. Carter wins the elections due to his strength in the South and East which secures him an insurmountable lead that Ford's near sweep of the West cannot overtake. Carter gets 342 electoral votes whereas Ford receives only 196. Moreover, Carter gains 28 states and the District of Columbia whereas Ford wins in only 22 states.

These scenarios reflected the most recent polls. Participants were instructed to imagine that the given scenario was true and, in addition, to imagine the winner's victory speech and the loser's concession of defeat. Thus the overall image was intended to be as plausible and vivid as possible.

After having imagined a particular outcome participants were asked to predict how they thought the election would actually turn out.

Result:

Participants' predictions corresponded to imagined scenarios: Participants having imagined Ford as the winner predicted predominantly Ford as the next president and vice versa.

Interpretation and criticism:

The result of this study is difficult to interpret since the two conditions are associated with different plausible scenarios as well as with imagining these different scenarios. Due to this confounding it is impossible to attribute the effect uniquely to the process of imagining the scenarios (cf. the discussion concerning confounding in Chapter 2.6.2).

The following study provides a better illustration of the relation between vivid imagination and availability by demonstrating an increased or decreased subjective probability of an event depending on the difficulty to generate an image.

Ex. 4-8: Availability and the difficulty to generate an image (Sherman, Cialdini, Schwartzman, and Reynolds, 1985):

The experiment was concerned with the possibility to become infected with the hypothetical disease *Hyposcenia-B* that was rampant on the campus of the Arizona State University.

The experiment comprised two factors:

1. *Instruction to imagine a scenario vs. no such instruction:*
   In the condition with instruction participants had to read the description of the symptoms of the disease, and then they should imagine to have been infected by the disease and to suffer from its reported symptoms for three weeks.
   In the condition without instruction participants had to read the description of the symptoms only.
2. *Concrete symptoms that are easy to imagine vs. abstract symptoms that are difficult to imagine:*
   In the condition with easy to imagine symptoms the latter were quite concrete: muscle aches, low energy level, and frequent severe headaches.
   In the condition with difficult to imagine symptoms the latter were abstract: a vague sense of disorientation, a malfunctioning nervous system, and an inflamed lever.

120 female students were assigned randomly to the four experimental conditions resulting from the combinations of the two factors.

Having read and, according to experimental condition, created an image of the symptoms participants had to rate the possibility of becoming infected by the disease on a scale from 1 (very probable) to 10 (quite improbable).

Results:

- ❑ Participants in the condition with imagination of easy to imagine symptoms provided lower ratings than participants without imagination (of easy to imagine symptoms): 5.25 vs. 6.20, indicating a higher subjective probability of getting infected.
- ❑ Participants in the condition with imagination of difficult to imagine symptoms provided *higher* ratings than participants without imagination (of difficult to imagine symptoms): 6.55 vs. 7.70, indicating a higher subjective probability of getting infected.
- ❑ The abstractness of the imagined symptoms had an effect in the conditions with the instruction to imagine: 5.25 vs. 7.70.
- ❑ The abstractness of the imagined symptoms had a slight effect only in the conditions without the instruction to imagine: 6.20 vs. 6.55.

Interpretation:

The experiment demonstrates the effect of the availability on the subjective probability to acquire a disease: With easy to imagine symptoms imaging results in greater subjective probability to get infected whereas in case of difficult to imagine symptoms the opposite effect was observed: Imagining abstract symptoms leads to a lower subjective probability compared to no imagining.

The illustrated effect of imagining on subjective probabilities of certain events may be relevant in daily life. Specifically it might explain the great proportion of annual insolvencies and other failures: If people already see themselves in their imagination as the new »Bill Gates«, »Steve Jobs« or, say, »Frank Sinatra« they are prone to underestimate the risks and the fact that only very view people make it to the top. The same factor might also be relevant in case of lotto players: People often imagine detailed scenarios of their future life as millionaires thereby ignoring the low chances (see also Section 5.4.2).

4.2.1.2 Vividness, Personal Experience, and Availability

One important factor determining the degree of availability concerns the vividness and personal involvement of the subject. The following example illustrates this aspect.

Ex. 4-9: A real-life anecdote (after Nisbett & Ross, 1980):

Urs Bürli intends to buy a new car. He decides to purchase a Swedish car: a Saab or a Volvo. Since he is unsure with respect to the relative robustness of both brands he studies the relevant journals and repair statistics. On the basis of this information that includes hundreds of cases he finally decides to buy a Saab.

On Sunday evening just before fixing the purchase he meets his friend Hans-Rudi at his regular's table in their pub and he talks with him about his decision. Hans-Rudi reacts with astonishment: »I have an acquaintance possessing a Saab. First the brakes were defect then the injection pump broke. Furthermore there were always troubles with the electronics. After 3 years he sold the car for peanuts«.

Due to the information from his friend Urs Burli decides to purchase a Volo.

This story illustrates nicely the selectivity of human information processing: On the basis of a single vivid report the evidence from hundreds of cases is ignored.

Generally, humans assign higher weight to information from personal experiences than to abstract statistical information. This also explains why Steven Spielberg's film Schindler's list had a much greater impact on peoples' perception of the atrocities of the Nazi regime than abstract statistics about how many millions of Jewish people were kil-

led by the regime. The reason is obvious: The film portrays realistically the fate of individual people thus enabling one to witness the concrete destinies of these persons.
A typical symptom of our liability to concrete and personal information is the vice to argue by means of examples and anecdotes that is also common in academic circles. For example, one can hear arguments like the following: »My grand farther smoked more than 40 cigarettes a day and lived nevertheless for over 90 years«.
Unfortunately, the only reasonable answer to statements like this is given only rarely:

»Ok, but what do you want to prove with your statement? Dozens of empirical studies have found a positive relation between smoking and cancer as well as between smoking and a shortened life time«.

Principle 4-1: *Examples and General Statements:*

Examples are useful to illustrate principles. However, they are inappropriate for supporting statements about general relationships.

The final example illustrates that personal involvement can result in biased judgments concerning the distribution of political power.

Ex. 4-10: Vividness and erroneous judgments of political conditions (Nisbett & Ross, 1980):

During the presidential elections in 1972 a great number of journalists participating the campaign of McGovern claimed steadfastly (and incorrectly) that the distance in votes between the latter and his opposite will not be greater than 10 percent despite the fact that all the surveys predicted a distance of at least 20 percent and despite their knowledge that that the polls have never erred by more than 3 percent.
The reason for this erroneous judgment is found in the fact that they witnessed with their own eyes the thousands of enthusiastic fans of McGovern.

A comparable overestimation of real political conditions is often found in groups articulating themselves forcefully and frequently in the public. The political influence of these groups is also commonly overestimated by a majority of people and also by the politicians.

Let us now turn to examples illustrating the functioning of the representativeness heuristic.

4.2.2 The Representativeness Heuristic and its Effects

One effect of the representativeness heuristic, the misconception of randomness and the resulting consequences, has already been demonstrated in the context of erroneous judgments of random sequences (cf. Section 2.3).

A classical study illustrating the functioning of the representativeness heuristic concerns the so called *conjunction fallacy*.

Ex. 4-11: The conjunction fallacy (Tversky and Kahneman, 1983):

Participants:

Three groups with different degrees of statistical sophistication:

1. Undergraduates without statistical knowledge;
2. Graduates with intermediate statistical knowledge;
3. Graduates with good statistical knowledge.

Stimuli and procedure:

1. Bill:

The participants received the following personality sketch and instruction:

Bill is 34 years old. He is intelligent, but unimaginative, compulsive, and generally lifeless. In school, he was strong in mathematics but weak in social studies and humanities.

Please rank the following statements according to their probability where »1« indicates »most probable« and »8« signifies »most improbable«.

Result:

Ranking	Statement	
(4.1)	Bill is a physician who plays poker for a hobby.	
(4.8)	Bill is an architect.	
(1.1)	Bill is an accountant.	(***B***)
(6.2)	Bill plays jazz for a hobby.	(***J***)
(5.7)	Bill surfs for a hobby.	
(5.3)	Bill is a reporter.	
(3.6)	Bill is an accountant who plays jazz for a hobby.	($\boldsymbol{B} \wedge \boldsymbol{J}$)
(5.4)	Bill climbs mountains for a hobby.	

2. Linda:

Personality sketch and instruction:

Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Please rank the following statements according to their probability where »1« indicates »most probable« and »8« signifies »most improbable«.

Result:

Ranking	Statement	
(5.2)	Linda is a teacher in elementary school.	
(3.3)	Linda works in a bookstore and takes Yoga classes.	
(2.1)	Linda is active in the feminist movement.	(***F***)
(3.1)	Linda is a psychiatric social worker.	
(5.4)	Linda is a member of the League of Women Voters.	
(6.2)	Linda is a bank teller.	(***B***)
(6.4)	Linda is an insurance salesperson.	
(4.1)	Linda is a bank teller and is active in the feminist movement.	($\boldsymbol{B \wedge F}$)

In both cases participants commit the so called conjunction fallacy: They judge the conjunctive event as being more probable than one of its components. In case of Bill the conjunctive event $\boldsymbol{B \wedge J}$ is judged as being more probable than the event ***J***. In case of Linda $\boldsymbol{B \wedge F}$ is assessed as more probable than ***B***.

The following table shows that the degree of statistical education has practically no influence on the result.

Tab. 4-1: *Percentage of rankings in which the conjunctive event is ranked prior to the single events as a function of statistical knowledge for the two scenarios.*

	Naive		Intermediate Level		High Level	
	Linda	Bill	Linda	Bill	Linda	Bill
	89%	93%	90%	86%	85%	83%
N	88	94	53	56	32	32

Interpretation:

The presented personal descriptions are highly representative for accountants and active feminists. However, they are untypical for jazz players and bank tellers. Consequently, participants rank the probability of conjunctive events (*Bill is an accountant who plays jazz for a hobby* and *Linda is a bank teller and is active in the feminist movement*) in between those of the two single events.

The probability of the events is thus estimated on the basis of the similarity to the general description of Bill and Linda and not due to the set theoretic relations between the sets involved as prescribed by the axioms of probability (cf. appendix).

A debriefing of participants revealed that, except for two persons, all of them accepted the statistical principle. In addition they revised their erroneous judgments.

A possible objection:

The following objection against the presented interpretation might be asserted (which is, according to my opinion, somewhat far-fetched):

Participants have interpreted the statement:

Linda is a bank teller and *Bill plays jazz for a hobby*

as follows

Linda is a bank teller who is not active in the feminist movement, and

Bill plays jazz for a hobby and he is not an account.

To exclude this possible objection an experiment was conducted in which only the untypical feature had to be evaluated (the typical feature was not presented).

In each case the conjunctive event was rates as more probable than the singular event making up the conjunctive one.

In a further study Tversky and Kahneman investigated the role of representativeness on predictions. In this case representativeness is interpreted as »Consequence *X* is more typical for cause *M*«.

Ex. 4-12: The conjunction fallacy and predictions (Tversky & Kahneman, 1982a, p.96):

The subsequent predictions concerning the year 1981 had to be ordered by participants according to their probability:

Results:

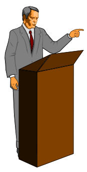
Rank	Statement
(1.5)	Reagan will cut federal support to local government. (***B***)
(3.3)	Reagan will provide support for unwed mothers. (***A***)
(2.7)	Reagan will increase the defense budget by less than 5%.
(2.9)	Reagan will provide federal support for unwed mothers and cut federal support for local government. ($\boldsymbol{A \wedge B}$)

The conjunction error was committed by 68 percent of the participants.

The presented examples give reason to the following warning:

Beware of detailed internally coherent and plausible scenarios (those concerning the future as well as those concerning the past).
More detailed scenarios appear as more plausible. However more detailed scenarios are less probable since each added detail reduces the probability of the scenario.

On the basis of these results Tversky und Kahneman (1983) conclude that people reason *intuitively*, that is, they base their judgments on typicality or on considerations of similarity. The axioms of mathematical probability, on the other hand, refer to *extensional* structures. Specifically the probability measure is defined on sets and relation between sets. Accordingly the probability assigned to the conjunction $A \wedge B$ of two events A and B, represented by the intersection $A \cap B$ (events are generally represented by sets), can never be smaller than the probability assigned to the two events since the intersection of two sets is always included in each of the two sets.
The conjunction fallacy demonstrates that people do not reason in the same way by taking relations between extensional structures into account as demanded by probability theory. However, it is possible to increase the tendency to reason extensionally by using absolute frequencies instead of probabilities or by increasing the saliency of the relations between events (cf. Chapter **Error! Reference source not found.**).



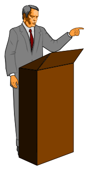
Comment 4-2: Conjunction fallacy and different conceptions of probability

As noted above in Chapter 4.1, the mathematical conception of probability is but one of a number of different views of probability. The *subjective view* that regards probability as *(rational) degree of belief* is not an extensional conception of probability i.e. subjective probabilities are not defined on sets.
However, the subjective view, like the other conceptions of probability, accepts the axioms of probability, and, by consequence, the proposition that a conjunctive event must be lower than (or equal to) the probabilities of each of the other events making up the conjunction.

Following to this review of the biasing effects of the availability and representativeness heuristic we next turn to problems of people in interpreting probability information.

4.3 Difficulties in Understanding Probability Information

The present section discusses misunderstandings and the ignorance of relevant probability information. Specifically, the following three topics are discussed:
1. Problems in handling conditional probabilities;
2. Ignoring base rate information;
3. Risk communication in medicine.



Comment 4-3: Over- and underweighting of probability information

A further aspect with respect to biases in handling probability information concerns the overweighting of small probabilities close to 0.0, and the underweighting of probabilities close to 1.0. This issue will be treated in Chapter 5.4.2.

4.3.1 Conditional Probabilities: Formal and Psychological Aspects

The concept of conditional probabilities and the operation of conditioning are of great importance in probabilistic judgments and reasoning. In the following, we first discuss important formal characteristics of the concept. Then we turn to typical errors concerning the understanding and use of conditional probabilities.

4.3.1.1 CONDITIONAL PROBABILITY AND ITS CHARACTERISTICS

Concept 4-1: *Conditional Probability:*

Given: Two random variables X and Y.

The conditional probability $P(X = x | Y = y)$ represents the probability that the random variable X takes on the value x given that the value of the random variable Y has the value y.

Explication:
1. A random variable is a variable that takes on different values with different probabilities.
2. $X = x$ denotes the specific *event* that X takes on the value x and $Y = y$ denotes the event that Y takes on the value y.

Interpretation:
1. Within the classical conception, the conditional probability refers to the probability that the variable X takes on the value x within the *sub-population* that is given by the fact that each member of this sub-population exhibits the value y with respect to the variable Y.

Thus the conditional distribution represents the probability that $X = x$ (The value of random variable *X* is *x*) not with respect to the whole population (where the values of *Y* can vary too) but with respect to the sub-population where the condition $Y = y$ holds for each of the members.

2. Within the subjective conception, the conditional distribution refers to the subjective probability of $X = x$, given that the subjects has the knowledge that the value of variable *Y* is *y*.

Representation of conditional probabilities by means of joint and marginal probabilities:
The conditional probability can be represented by the following formula:

$$P\left(X = x \middle| Y = y\right) = \frac{P\left(X = x \wedge Y = y\right)}{P\left(Y = y\right)},$$

Interpretation (objective conception):
The formula representing the conditional probability in terms of the joint and marginal probabilities states that the conditional probability of the event $X = x$ within the sub-population that is characterized by the condition $Y = y$ results by counting the number of entities with both conditions being satisfied (i.e. $X = x$ and $Y = y$) and dividing by the size of sub-population (i.e. the number of units meeting the condition $Y = y$).

Interpretation (subjective conception):
The formula representing the conditional probability in terms of the joint and marginal probabilities states that the conditional probability of the event $X = x$ given that the event $Y = y$ is known to have occurred corresponds to the probability of the joint event $X = x$ and $Y = y$ divided by the probability that event $Y = y$ occurs.

The following example demonstrates the central aspects of conditional probabilities.

Ex. 4-13: Conditional probability

Given:

Two random variables *X* and *Y* that indicate smoking and lung cancer, specifically:

$$X = \begin{cases} S & \Leftrightarrow \quad \text{subject is smoking} \\ \overline{S} & \Leftrightarrow \quad \text{subject is not smoking} \end{cases}$$

$$Y = \begin{cases} L & \Leftrightarrow \quad \text{subject has lung cancer} \\ \overline{L} & \Leftrightarrow \quad \text{subject has not lung cancer} \end{cases}$$

Figure 4-2 shows a population of smokers and non-smokers having either lung cancer or not. Each point represents a subject.

The whole population can be partitioned into four groups:

$S \wedge L$: Smokers that have lung cancer.
$S \wedge \bar{L}$: Smokers that don't have lung cancer.
$\bar{S} \wedge L$: Non-smokers that have lung cancer.
$\bar{S} \wedge \bar{L}$: Non-smokers that don't have lung cancer.

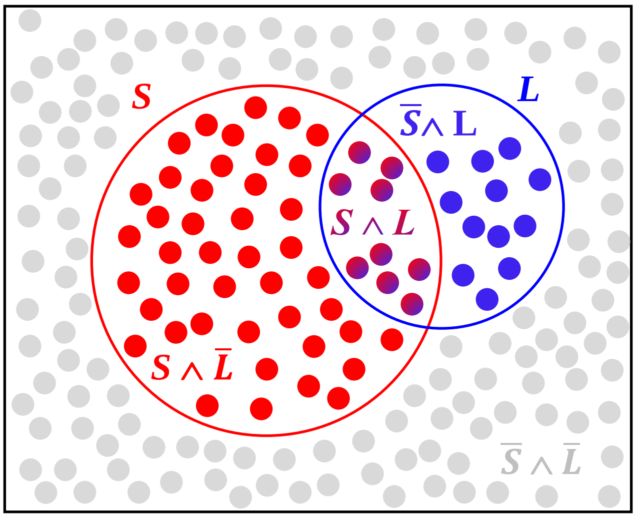


Figure 4-2: *Venn diagram illustrating the central aspects of conditional probabilities (S = smokers, L = lung cancer).*

Comments:

1. The four groups *partition* the population. This means that the groups are *exclusive* (each subject in the population belongs to one of the four groups only) and *exhaustive* (the four groups make up the whole population).
2. Note that there exist further partitions of the population, like the partititon of smokers versus non-smokers or the pratitions consisting of people with and without lung cancer. These partitions are not as fine grained as the partition on the basis of the joint events.

Wanted:

The conditional probability $P(X = S | Y = L)$ of a subject being a smoker given that she has lung cancer.

Solution:
Applying the equation representing the conditional probability in terms of the joint and the marginal probability:

$$P\left(X=S\middle|Y=L\right)=\frac{P\left(X=S\wedge Y=L\right)}{P\left(Y=L\right)},$$

enables one to retrieve the solution directly from Figure 4-2:

$$P\left(X=S\middle|Y=L\right)=\frac{\#\left(X=S\wedge Y=L\right)/N}{\#\left(Y=L\right)/N}$$

$$=\frac{\#\left(X=S\wedge Y=L\right)}{\#\left(Y=L\right)}=\frac{9}{21}=\underline{\underline{\frac{3}{7}}}$$

N denotes the whole number of subjects in the population and the symbol # indicates the number of cases.

Interpretation:
The conditional probability represents the proportion of smokers within the sub-population made up be the subjects with lung cancer.

Comments:

1. The conditional probability is given by the number of points being both blue and red, divided the sum of the number of points being both blue and red as well as the number of points being blue only (this latter sum is given by all the points within the blue circle).
2. The conditional probability $P\left(X=S\middle|Y=L\right)$ differs from the *inverse conditional probability* $P\left(Y=L\middle|X=S\right)$. The latter is given by the number of blue and red points divided by the number of points within the red circle. Thus, the inverse probability $P\left(Y=L\middle|X=S\right)$ is considerably smaller than $P\left(X=S\middle|Y=L\right)$.
3. Consequently, conditional probabilities are usually *not symmetric*, i.e., the conditional probability of $X = x$ given $Y = y$ is not the same as the inverse conditional probability of $Y = y$, given $X = x$. This occurs only if the two sub-populations involved are of the same size. In the present case the sub-population of smokers has to be of the same size as the sub-population of people with lung cancer (which is not the case).

4. The conditional probability $P(X = x|Y = y)$ is greater than the inverse probability $P(Y = y|X = x)$ if the sub-population given by $Y = y$ is smaller than the sub-population given by $X = x$.
5. $P(X = \bar{S}|Y = L) = 1 - P(X = S|Y = L)$.
   This equality follows from the fact that the probabilities of the different values within a population and subpopulation, respectively, must sum to 1.
6. In general, $P(X = S|Y = L) \neq P(X = S|Y = \bar{L})$ since the value of a random variable need not have the same probability in different populations.

Previously to discussing further aspects of the concept some notational conventions have to be introduced.

Notation 4-1:

1. Instead of using the somewhat cumbersome notation $P(X = S|Y = L)$ the abbreviation $P(S|L)$ will be used [In general: $P(x|y)$ instead of $P(X = x|Y = y)$].
2. The symbol $P(X|Y = y)$ and $P(X|y)$, respectively, refers to the distribution of the values of variable *X* given the specific value *y* of *Y*.
   In the discrete case the symbol can be conceived of as denoting a list that contains the probabilities of each of the values of the random variable *X* for the sub-population given by $Y = y$.
3. The symbol $P(X|Y)$ refers to the distribution of the values of *X* for every value of variable *Y*.
   In the discrete case the symbol can be conceived of as denoting a table that contains the probabilities of each of the values of the random variable *X* for each value of variable *Y*.

Ex. 4-14: Notation of conditional probability

Given: The two variables:

$$X = \begin{cases} S & \Leftrightarrow \text{ subject is smoking} \\ \bar{S} & \Leftrightarrow \text{ subject is not smoking} \end{cases}$$

$$Y = \begin{cases} L & \Leftrightarrow \text{ subject has lung cancer} \\ \bar{L} & \Leftrightarrow \text{ subject has not lung cancer} \end{cases}$$

The symbol $P(X|Y)$ refers to the following table of conditional probabilities:

X	Y: L	$\bar{L}$
S	$P(X=S\|Y=L)$	$P(X=S\|Y=\bar{L})$
$\bar{S}$	$P(X=\bar{S}\|Y=L)$	$P(X=\bar{S}\|Y=\bar{L})$

The symbol $P(X|Y=L)$ refers to the column labeled L whereas $P(X=S|Y)$ refers to the row with $X=S$.

In the general case, with variable X assuming n possible values $x_1, x_2, \ldots, x_n$ and Y assuming m possible values $y_1, y_2, \ldots, y_m$, $P(X|Y)$ refers to the $m \times n$ table of conditional probabilities and the symbols $P(X=x_i|Y)$ and $P(X|Y=y_j)$ $(i=1,2,\ldots,n)$, $(j=1,2,\ldots,m)$ refer to the respective rows and columns.

X	Y: y_1	y_2	$\cdots$	y_m
x_1	$P(X=x_1\|Y=y_1)$	$P(X=x_1\|Y=y_2)$	$\cdots$	$P(X=x_1\|Y=y_m)$
x_2	$P(X=x_2\|Y=y_1)$	$P(X=x_2\|Y=y_2)$	$\cdots$	$P(X=x_2\|Y=y_m)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	$P(X=x_n\|Y=y_1)$	$P(X=x_n\|Y=y_2)$	$\cdots$	$P(X=x_n\|Y=y_m)$

In many situations, the values of one random variable H represent different hypotheses $H_1, H_2, \ldots, H_n$ that partition the so called *hypothesis space*. Typical cases are hypotheses concerning the presence or absence of a disease, the guiltiness of a defendant, the toxicality of a substance, etc.

The second random variable E represents empirical evidence e.g. the outcome of a diagnostic test, the testimony of a witness or the outcome of animal studies concerning a possibly toxical substance.

Figure 4-3 depicts the hypothesis space (the area within the rectangle) and its partitioning according to the possible combinations of the outcomes of the two variables H and E. The whole hypothesis space is assumed to have size 1 and the sizes of the different regions represent the probabilities of the associated events.

4.3.1.2 Revising Probabilities by Means of Conditioning

One important issues in probabilistic reasoning concerns the problem of how the probability of a hypothesis changes in the light of the given evidence. This problem can be restated using conditional probabilities:

Given the probability $P(H)$ of hypothesis H, the conditional probability $P(H|E)$ represents the probability of H given evidence E. Thus, by teking evidence E into account the probability $P(H)$ can be replaced by (or updated to) $P(H|E)$. The latter may be computed using the definition of the conditional probabilities in terms of the joint and the marginal probabilities:

$$P(H|E) = \frac{P(H \wedge E)}{P(E)}.$$

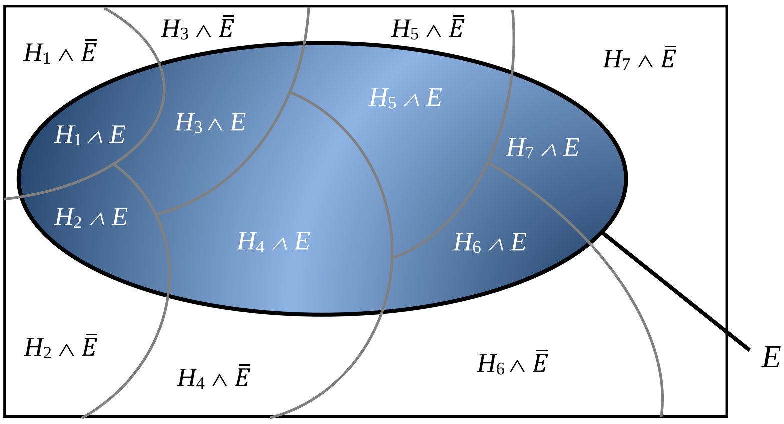
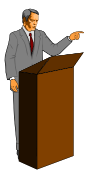


Figure 4-3: *A hypothesis space (area within the rectangle) made up by all possible combinations of hypotheses and outcomes. The size of the different areas represents the probability of the associated values of the variables and combination of variables, respectively.*
The set of hypotheses $H_1, H_2, \ldots, H_7$ are dsjoint and exhaustive and thus partition the whole space. Similarly E and $\bar{E}$ partition the whole space. Consequtly the joint events $H_1 \wedge E, \ldots, H_7 \wedge E, H_1 \wedge \bar{E}, \ldots, H_7 \wedge \bar{E}$ partition the hypothesis space too.

The operation of updating probabilities by computing the conditional probability is also called *conditioning* on the event E. The original probability $P(H)$ of hypothesis H is called the *prior probability*, and the revised probability $P(H|E)$ that takes the evidence into account is termed the *posterior probability*. This terminology is used in the context of Bayes theorem (cf. Notation 4-4, on page 169).



Comment 4-4: Bayes formula

The equation,

$$P\left(H|E\right)=\frac{P\left(H\wedge E\right)}{P\left(E\right)},$$

represents an elementary version of Bayes theorem (for more details cf. Section 4.4.1.5).

Ex. 4-15: Updating of probabilities

Consider, once again, Figure 4-3 and specifically hypothesis H_3. The probability $P\left(H_3\right)$ of H_3 is given by the associated area consisting of the sum of the two areas labeled $H_3 \wedge E$ and $H_3 \wedge \overline{E}$ divided by the whole area of the rectangle (which is assumed to be one).

Consequently, the conditional probability $P\left(H_3|E\right)$ is given by:

$$P\left(H_3|E\right)=\frac{P\left(H_3\wedge E\right)}{P\left(E\right)}=\frac{\text{Area of }\left(H_3\wedge E\right)}{\text{Area of }E},$$

The area of E corresponds to the area of the black shaded ellipse. Obviously, $P\left(H_3|E\right)$ is greater than $P\left(H_3\right)$. Consequently, evidence E is in favor of H_3, thus increasing its probability.

4.3.1.3 Conditional Probabilities with more than two Variables.

The previous discussion involved two variables only. However, the concept and definition can be generalized to more than two variables:

1. The sub-population, we are interested in, may be characterized by more than a single condition. Specifically the sub-population may be characterized by the joint event: $Y_1 = y_1 \wedge Y_2 = y_2 \wedge \ldots \wedge Y_m = y_m$ with the given joint event being met by all members of the sub-population. For example, one might be interests in the probability of a subject having lung cancer within the sub-population consisting of male smokers and of age greater than or equal to 50 years:

   $$Y_1 = \text{male} \wedge Y_2 = \text{smoker} \wedge Y_3 = \text{age} \geq 50.$$

   The requested conditional probability of having lung cancer within the indicated target population may thus be denoted by the symbol:

   $$P\left(X = \text{lung cancer}|Y_1 = \text{male} \wedge Y_2 = \text{smoker} \wedge Y_3 = \text{age} \geq 50\right)$$

   [or, $P\left(\text{lung cancer}|\text{male} \wedge \text{smoker} \wedge \text{age} \geq 50\right)$, for short].

2. In addition the event whose probability we are interested in may be characterized by the joint values of two and more random variables: $X_1 = x_1 \wedge X_2 = x_2 \wedge \ldots \wedge X_n = x_n$. For example we may by interests in the probability of people suffering from lung cancer and a high

blood pressure for male smokers of age greater than or equal to 50 years. This probability may denoted by:

$P\left(X_1 = \text{lung cancer} \wedge X_2 = \text{high bood pressure} \middle| Y_1 = \text{male} \wedge Y_2 = \text{smoker} \wedge Y_3 = \text{age} \geq 50\right)$

or, more compact,

$P\left(\text{lung cancer} \wedge \text{high bood pressure} \middle| \text{male} \wedge \text{smoker} \wedge \text{age} \geq 50\right)$.

The inclusion of additional variables neither changes the basic interpretation of conditional probabilities nor their definition in terms of joint and marginal probabilities:

$$P\left(X_1 = x_1 \wedge X_2 = x_2 \wedge \ldots \wedge X_n = x_n \middle| Y_1 = y_1 \wedge Y_2 = y_2 \wedge \ldots \wedge Y_m = y_m\right)$$
$$= \frac{P\left(X_1 = x_1 \wedge X_2 = x_2 \wedge \ldots \wedge X_n = x_n \wedge Y_1 = y_1 \wedge Y_2 = y_2 \wedge \ldots \wedge Y_m = y_m\right)}{P\left(Y_1 = y_1 \wedge Y_2 = y_2 \wedge \ldots \wedge Y_m = y_m\right)} \cdot$$

Referring to our example, the equation for the conditional probability looks like this:

$$P\left(\text{lung cancer} \wedge \text{high blood pressure} \middle| \text{male} \wedge \text{smoker} \wedge \text{age} \geq 50\right)$$
$$= \frac{P\left(\text{lung cancer} \wedge \text{high blood pressure} \wedge \text{male} \wedge \text{smoker} \wedge \text{age} \geq 50\right)}{P\left(\text{male} \wedge \text{smoker} \wedge \text{age} \geq 50\right)}$$

Thus, the conditional probability of the joint event of suffering from lung cancer and a high bood pressure within the population of male smokers with age greater or equal to 50 is given by the number of people having each of the features divided by the number of people within the population (of male smokers of age ≥ 50).

Ex. 4-16: Conditional probabilities with more than two variables

Given: The three variables:

$$X = \begin{cases} S & \Leftrightarrow \quad \text{subject is smoking} \\ \overline{S} & \Leftrightarrow \quad \text{subject is not smoking} \end{cases}$$

$$Y = \begin{cases} L & \Leftrightarrow \quad \text{subject has lung cancer} \\ \overline{L} & \Leftrightarrow \quad \text{subject has not lung cancer} \end{cases}$$

$$Z = \begin{cases} A & \Leftrightarrow \quad \text{subject of age} \geq 50 \\ \overline{A} & \Leftrightarrow \quad \text{subject of age} < 50 \end{cases}$$

Figure 4-4 depicts the partitioning of the sampling space and population, respectively (represented by the rectangle) into eight regions due to the combination of the presence and absence of the three characteristics (= the combination of the values of the three random variables).

On the basis of the probabilities of the joint events all sorts of conditional probabilities can be computed, for example:

$$P(L|S\wedge A)=\frac{P(S\wedge L\wedge A)}{P(S\wedge A)}=\frac{\#(S\wedge L\wedge A)}{\#(S\wedge A)}$$

$$=\frac{\#(L\wedge S\wedge A)}{\#(S\wedge L\wedge A)+\ \#(S\wedge \bar{L}\wedge A)}=\frac{7}{7+3}=\underline{\underline{\frac{7}{10}}}$$

$$P(S\wedge A|L)=\frac{P(S\wedge L\wedge A)}{P(L)}=\frac{\#(S\wedge L\wedge A)}{\#(L)}$$

$$=\frac{\#(L\wedge S\wedge A)}{\#(S\wedge L\wedge A)+\ \#(S\wedge L\wedge \bar{A})+\#(\bar{S}\wedge L\wedge A)+\ \#(\bar{S}\wedge \bar{L}\wedge \bar{A})}$$

$$=\frac{7}{7+10+4+11}=\underline{\underline{\frac{7}{32}}}$$

Comment:
This illustrates, once again, the asymmetry of conditional probabilities.

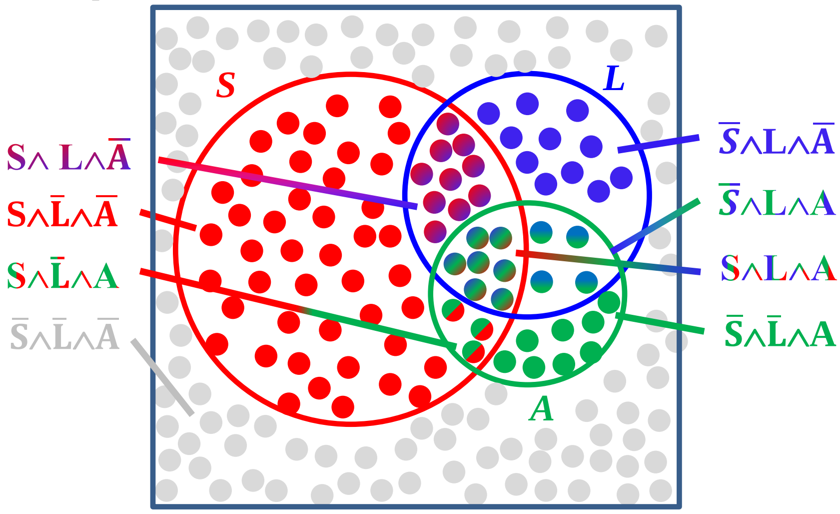


Figure 4-4: *Venn diagram illustrating the partioning of the population into eight regions in case of three variables with two values each (S = smokers, L = lung cancer, A = age≥50).*

Notation 4-2:

In the following the symbol $\wedge$ (representing the logical *and*) will be replaced by a comma. Thus, the symbol:

$P(X_1=x_1\wedge X_2=x_2\wedge\ldots\wedge X_n=x_n|Y_1=y_1\wedge Y_2=y_2\wedge\ldots\wedge Y_m=y_m)$

will be replaced by the symbol:

$P(X_1=x_1,X_2=x_2,\ldots,X_n=x_n|Y_1=y_1,Y_2=y_2,\ldots,Y_m=y_m)$ or

$P(x_1,x_2,\ldots,x_n|y_1,y_2,\ldots,y_m)$.

Similarly, the symbol $P(X_1\wedge X_2\wedge\ldots\wedge X_n|Y_1\wedge Y_2\wedge\ldots\wedge Y_m)$ will be replaced by $P(X_1,X_2,\ldots,X_n|Y_1,Y_2,\ldots,Y_m)$.

Notation 4-3:

The conventions of Notation 4-1 (p. 127) will be accommodated to the case with more than two variables. Thus, the symbols:

$P\left(X \middle| Y_1 = y_1, Y_2 = y_2, \ldots, Y_m = y_m\right)$ and $P\left(X \middle| y_1, y_2, \ldots, y_m\right)$,

respectively, refer to the distribution of the variable X within the population that is characterized by the combination of values $y_1, y_2, \ldots, y_m$. This distribution represents, in the discrete case, the probability that X assumes a given value x, for each value x of variable X. Thus, the symbol refers to a whole list of probabilities (cf. Ex. 4-17).

More general, $P\left(X_1, X_2, \ldots, X_n \middle| Y_1, Y_2, \ldots, Y_m\right)$ represents, in the discrete case, the probabilities of all combination of values of the variables $X_1, X_2, \ldots, X_n$ given all possible values of the variables $Y_1, Y_2, \ldots Y_m$. Thus, the symbol refers to a whole table of probabilities (cf. Ex. 4-17).

Ex. 4-17: Notation of conditional probabilities with more than two variables

Given: The four variables:

$$X = \begin{cases} 1 & \Leftrightarrow \quad \text{Feature } X \text{ is present} \\ 0 & \Leftrightarrow \quad \text{Feature } X \text{ is absent} \end{cases}$$

$$Y = \begin{cases} 1 & \Leftrightarrow \quad \text{Feature } Y \text{ is present} \\ 0 & \Leftrightarrow \quad \text{Feature } Y \text{ is absent} \end{cases}$$

$$Z = \begin{cases} 1 & \Leftrightarrow \quad \text{Feature } Z \text{ is present} \\ 0 & \Leftrightarrow \quad \text{Feature } Z \text{ is absent} \end{cases}$$

$$W = \begin{cases} 1 & \Leftrightarrow \quad \text{Feature } W \text{ is present} \\ 0 & \Leftrightarrow \quad \text{Feature } W \text{ is absent} \end{cases}$$

The symbol $P\left(X, Y \middle| Z, W\right)$ refers to the table with the single conditional probabilities as entries:

		$W = 1$		$W = 0$	
X	Y	$Z = 1$	$Z = 0$	$Z = 1$	$Z = 0$
1	1	$P(1,1 \mid 1,1)$	$P(1,1 \mid 0,1)$	$P(1,1 \mid 1,0)$	$P(1,1 \mid 0,0)$
	0	$P(1,0 \mid 1,1)$	$P(1,0 \mid 0,1)$	$P(1,0 \mid 1,0)$	$P(1,0 \mid 0,0)$
0	1	$P(0,1 \mid 1,1)$	$P(0,1 \mid 0,1)$	$P(0,1 \mid 1,0)$	$P(0,1 \mid 0,0)$
	0	$P(0,0 \mid 1,1)$	$P(0,0 \mid 0,1)$	$P(0,0 \mid 1,0)$	$P(0,0 \mid 0,0)$

The symbol $P(X,Y|Z=1,W=1)$ refers to the column headed $Z=1$ and $W=1$.

Having introduced the most important characteristics of conditional probabilities there remains one important concept that can be specified (and elucidated) by means of conditional probabilities.

4.3.1.4 CONDITIONAL PROBABILITIES AND CONCEPTS OF STOCHASTIC INDEPENDENCE

The concept of conditional probability is used for characterizing the concept of *(conditional) stochastic independence* Let us start with a formal definition of the construct of stochastic independence.

Concept 4-2: *Stochastic (statistical) independence:*

Given: Two random variables *X* and *Y*.

X and *Y* are *stochastically independent*, if and only if (iff) the conditional distribution of *X* given any value *y* of *Y* corresponds to the marginal distribution of *X*, in symbols:

$P(X|Y=y)=P(X)$, for all values *y* of *Y*

or, equivalently:

$P(Y|X=x)=P(Y)$, for all values *x* of *X*.

Due to the definition of conditional probabilities the two specifications are equivalent to the following one:
X and *Y* are *stochastically independent*, iff the joint distribution of *X* and *Y* conforms to the product of the marginal distributions of *X* and *Y*, in symbols:

$$P(X,Y)=P(X)\cdot P(Y)$$

The principle message of the definition of stochastic independence consists in the assertion that the distribution of the random variable *X* stays the same independently of which value the random variable *Y* assumes (and vice versa). By consequence, a consideration of the distribution of *X* need not take into account which value *y* that variable *Y* has assumed. Consequently, *Y* is completely irrelevant for determining the probability distribution of *X*.
The following example demonstrates the usefulness of the concept.

Ex. 4-18: Stochastic independence

Given:

Variable *T* represents a program for training dyslexic. It comprises two values:

$t \Leftrightarrow$ training

$\bar{t} \Leftrightarrow$ no training

Variable I represents the presence or absence of an improvement:

$i \quad \Leftrightarrow \quad$ improvement

$\bar{i} \quad \Leftrightarrow \quad$ no improvement

If it is the case that $P\left(I \middle| T=t\right) = P\left(I \middle| T=\bar{t}\right) = P\left(I\right)$, i.e. both variables are stochastically independent, then, obviously, the program has no effect.

The second important independence concept concerns conditional stochastic independence.

Concept 4-3: *Conditional stochastic (statistical) independence:*

Given: Three random variables *X*, *Y* and *Z*.

X and *Y* are *conditionally stochastic independent given Z*, iff the conditional distribution of *X* given *Y* and *Z* corresponds to the conditional distribution of *X* given *Z* alone, in symbols:

$P\left(X \middle| Y, Z=z\right) = P\left(X \middle| Z=z\right)$, for each value *z* of *Z*,

or, equivalently:

$P\left(Y \middle| X, Z=z\right) = P\left(Y \middle| Z=z\right)$, for each value *z* of *Z*,

This corresponds to the equivalent specification:
X and *Y* are stochastically independent iff the joint distribution of *X* and *Y*, given *Z* is identical to the product of the conditional distribution of *X* given *Z* and *Y* given *Z*, in symbols:

$P\left(X, Y \middle| Z\right) = P\left(X \middle| Z\right) \cdot P\left(Y \middle| Z\right)$

Conditional stochastic independence of *X* and *Y* given *Z* may be interpreted as follows:
For each fixed value of variable Z ($Z=z$) variable Y does not provide any information about variable X and vice versa.

Within the realm of psychology the concept of sonditional stochastic independence is important for, at least, two branches of reseach: Mediation research and psychometrics. Here are two examples:

Ex. 4-19: Mediation

Given:

A simple causal mediation model involving three variables:

S_E = Couple external stress (e.g. stress due to work conditions).

S_I = Couple internal stress: Stress that evolves during interactions with the partner.

S_P = Satisfaction with the partnership.

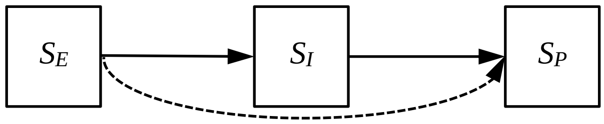


Figure 4-5: *Mediator model describing the effect of partner external stress on the satisfaction with the partnership.*

Figure 4-5 depicts the assumed causal relationships between the three variables. The model assumes that external stress has a direct influence on the partner internal stress which in turn has a direct influence on the statisfaction with the partnership. The dashed arrow indicates that there might also be a direct influence of S_E on S_P.

Assume, for the moment, that there is no direct influence of S_E on S_P. This situation is called *total mediation*. In this case, knowing the value of variable S_I makes the more distant variable S_E completely irrelevant with respect to the distribution of values of variable S_P. Thus, we have a situation of conditional independence between S_E and S_P given S_I:

$$P\left(S_P \middle| S_E, S_I\right) = P\left(S_P \middle| S_I\right)$$

The presence of conditional independence is easy to understand since, in case of total mediation, the causal effect of S_E on S_P is completely mediated by variable S_I. By consequence, S_E exerts an influence on S_P only by exerting an influence on the values of S_I, and only this value determines which value variabel S_P will take on. Thus, as soon as the value of S_I is known, S_E does not provide any further information about the value of S_P.

The situation is different in case of *partial mediation* where variable S_E exerts also a direct effect on S_P, additionally to the indirect effect via the mediator variable S_I (cf. the dashed arrow in Figure 4-5). In this case, conditional stochastic independence does no longer hold.

Ex. 4-20: The structure of psychometric models

Psychometric models consist of latent variables representing latent traits, latent states, etc. and a set of observed measurement (cf. *Figure* 2-5 on page 47).

It is assumed that the stochastic dependence between the observed variables is due to the fact that the measures are influenced by one or more common underlying traits.

Fixing the value of the underlying common traits makes the measures independent, i.e. within each sub-population with all units having the same values of the underlying traits the different measurements are independent.
Consequently, given the values of the common latent traits the measurements are independent or, expressed differently, the measurements are conditionally independent given the latent traits.

As already noted, independence does not imply conditional independence. But what about the reverse direction, does independence between two variables provide conditional independence between the two variables, given a third one? The answer to this question is *no*, as the following example demonstrates.

Ex. 4-21: Independence and conditional independence

In a psychiatric hospital two potential patients, Donald and Michael, spend the night in the same room.

- *Donald* exhibits a disposition to physically assault other people.
- *Michael* exhibits a disposition to paranoid hallucinations.

The next morning *Michael* complains himself of having been assaulted by *Donald* last night.
Consider the following three events:

A = Donald has assaulted Michael last night.
H = Michael had a paranoid attack last night thus having only hallucinated to have been assaulted by Donald.
R = Report of Michael of having been attacked by Donald.

Now, obviously A and H are independent events since there is no sensible relation between the tendency of Donald to assault Michael and the disposition of the latter to get a paranoid attack: $P(A|H) = P(A)$ and $P(H|A) = P(H)$, respectively.
However, given Michael's report R of having been attacked the two events are no longer independent since knowledge that Michael had a paranoid hallucination last night reduces my personal probability that he had been assaulted by Donald.
Similarly, knowing that Donald has assaulted Michael last night reduces my subjective probability of the latter having had a paranoid attack.
Thus: $P(A|H,R) < P(A|R)$ and $P(H|A,R) < P(H|R)$.

Ex. 4-21 illustrates a general causal situation that is of great importance in legal trials: Let C_1 and C_2 denote two potential independent causes that are able to elicit an effect E (In Ex. 4-21, A and H correspond to the two independent causes and R represents the effect).

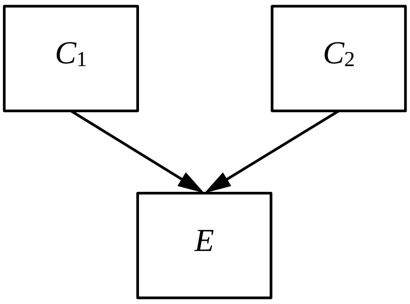


Figure 4-6: *Causal model with two independent causes C_1 and C_2 exerting a causal effect on a third variable E.*

This model plays a prominent role in legal trials for the following reason: Assume that a defendant is accused to have delivered a product that is known to possibly do harm to peoples' health (for example it is known to possibly elicit cancer). The aggrieved person thus demands a financial compensation. If the culprit is able to demonstrate that an alternative cause might be responsible for the harm (e.g. the harmed person was a heavy smoker) than this might advantageous to him since it reduces the probability that the harm had really been caused by his product and not by the alternative cause.
This terminates our discussuion of the formal aspects of conditional probability. In Chapter 4.4 we shall take up the issue again by considering in greater detail the basic probabilistic operations. We next turn to psychological aspects. Specifically, problems of the interpretation and use of conditional probabilities are considered.

4.3.1.5 The Inversion Error: Confusing Conditional Probabilities with their Inverses

As explicated above, in general the conditional probability of X give Y is not the same as the conditional probability of Y given X:

$$P\left(X|Y\right) \neq P\left(Y|X\right) \text{ [in general].}$$

However, lay persons sometimes confuse these two types of probabilities which may have quite serious consequences.

Ex. 4-22: Prosecutor's fallacy (Blitzstein & Hwang, 2015)

In 1998, Sally Clark was tried for murder after two of her sons died shortly after their birth.
During the trial, an expert witness for the prosecution testified that the probability of a newborn dying of sudden infant death syndrome (SIDS) was 1/8500. So the probability of two death due to SIDS in one family is $(1/8500)^2$ or about one in 73 million.

Thererfore, he continued, the probability of Clark's innocence was one 73 million.

Assessment of the expert's line of reasoning:

The expert commits two grave errors:

(a) He treats the two events (deaths) as being independent. However, it is highly probable that the sudden death is due to genetic or other family-specific risk factors.

(b) The expert commits the inversion error. He assumes that the probability of a crime given two sudden deaths is the same as the probability of two sudden deaths given a crime.
The number mentioned by the expert $(1/8500)^2$ is concerned with the later probability. The former is certainly much lower due to probability of a mother killing her two sons (This will be detailed in Section 4.4.1 where Bayes theorem is treated).

Sadly, Sally Clark was condemned of murder and spent three year in jail, partly due to the wrongheaded testimony of the expert, before her conviction was overturned. She died about one year after her release.

The problem of confusing conditional probabilities is frequently observed in case of diagnostic problems. Specificically, the probability of a disease given a positive diagnosis is confused with the probability of a positive diagnosis given a disease. Let us take a closer look at this problem.

Method 4-1: *Assessing the quality of a diagnostic instrument:*

Performing a diagnosis can be conceived of as probabilitistic decision problem. For the sake of concreteness, let us conceive of a medical diagnostic problems with the following events being involved:

D = disease present

$\bar{D}$ = disease absent

+ = positive diagnosis (indicating presence of the disease)

− = negative diagnosis (indicating absence of the disease)

The evaluation of a diagnostic instrument is based on the conditional probabilities shown in Tab. 4-2.

Tab. 4-2 *Conditional probabilities of a diagnostic decision problem.*

	Diagnosis		
Disease	+	–	Σ
D	$P(+\mid\mathrm{D})$	$P(-\mid\mathrm{D})$	1.0
$\bar{\mathrm{D}}$	$P(+\mid\bar{\mathrm{D}})$	$P(-\mid\bar{\mathrm{D}})$	1.0

In diagnostic contexts the shown conditional probabilities are given specific names:

- $P(+\mid\mathrm{D})$, the probability of a positive diagnosis, given the presence of the disease, is called the *sensitivity* of the diagnostic instrument. It corresponds to the *hit rate* of the instrument.
- $P(-\mid\bar{\mathrm{D}})$, the probability of a negative diagnosis, given the absence of the disease, is called the *specificity*. It corresponds to the *rate of correct rejection*.

The other cells are also given names that vary from context to context:

- $P(-\mid\mathrm{D})$ is called the *rate of false rejection*, or *Typ I* and *α error*, respectively.
- $P(+\mid\bar{\mathrm{D}})$ is called the *false alarm rate*, or *Typ II* and *β error*, respectively.

Note that the probabilities in each row sum to 1. Consequently, only one conditional probability in each row is required for a complete specification of the quality of the diagnostic instrument. Thus, knowing the sensitivity and specifity provides the complete information: The higher the sensitivity and the specificity the better the diagnostic instrument.

Unforunately many physician confuse the sensitivity $P(+\mid\mathrm{D})$ with the inverse probability $P(\mathrm{D}\mid+)$ that is the relevant one for the clients (cf. Casscells, Schönberger, & Grayboys, 1978; Eddy, 1982). Usually, the latter is much lower than the former, as demonstrated by the following example.

Ex. 4-23: Confusing sensitivity and prädictive posterior probability

Given: The following data:

- Prevalence of a disease D in the population: 0.3%
- Sensitivity of the diagnosis: 90% [$P(+\mid\mathrm{D}) = .90$]

- Specificity of a diagnosis: 97% [$P\left(-\middle|\bar{\mathrm{D}}\right)=.97$]

Question:
What is the probability that a randomly chosen member of the population with a positive diagnosis has in fact the disease? [$P\left(\mathrm{D}\middle|+\right)=?$]

Answer:
$P\left(\mathrm{D}\middle|+\right)=.083$. Thus the chances are only about 8.3% that a randomly chosen member of the population who received a positive result has actually the disease.

Interpretation:
The low probability is due to the fact that the prevalence of the disease in the population is low, and, by consequence, most positive outcomes of the diagnosis are false alarms.

Comment:
The result can be computed by means of Bayes theorem that will be discussed in Section 4.4.1.

4.3.1.6 THE NON-MONOTONICITY OF PROBABILITIES AND THE PROBLEM OF THE PROPER REFERENCE CLASS

One important aspect with respect to conditional probabilities with important implications for probabilistic reasoning concerns the following property:

Principle 4-2: *Non-monotonicity of conditional probabilities:*

Conditioning on an additional event can result in a reversal of the relationships between probabilities.

By consequence, it is possible that the following inequalities are true at the same time:

$P\left(X\middle|Y\right)>P\left(X\middle|Z\right)$,

$P\left(X\middle|Y,U\right)<P\left(X\middle|Z,U\right)$ and

$P\left(X\middle|Y,\bar{U}\right)<P\left(X\middle|Z,\bar{U}\right)$

The letters in the inequalities denote the presence and absence, respectively, of different events.

The non-monotonicity of conditional probabilities is at the heart of Simpson's paradox discussed in Section 2.6.3.

Ex. 4-24: Non-monotone probabilities and Simpson's paradox:

Consider once again Ex. 2-28 on page 63, concerning death sentences for Black and White delinquents in Florida.
Let:

D = Death sentence.

B_D = Black delinquent.

W_D = White delinquent.

B_V = Black victim.

W_V = White victim.

Obviously it is true that:

$P\left(D \mid W_D\right) > P\left(D \mid B_D\right)$, but

$P\left(D \mid W_D, B_V\right) < \left(D \mid B_D, B_V\right)$ and

$P\left(D \mid W_D, W_V\right) < \left(D \mid B_D, W_V\right)$

Thus, if only the color of the delinquent is taken into account, the probability of a death sentence is higher for Black than for White delinquents.
If, however, the color of the victim is also taken into account, the probability of the death sentence is higher for Black than for White delinquents, irrespectively of whether the victim is Black or White.

The lack of monotonicity of probabilities distinguishes probabilistic reasoning from deductive inference. In the latter case the validity of an inference can be assessed by means of pure local features, that is, without recurring to information other than those involved in the inference. Specifically, with deductive inference, if the premises are true and the inference scheme is correct then the conclusion must be true as well. The addition of further premises does not change the correctness of the conclusion.
The lacking monotonicity of probabilities has practical consequences in the context of inductive reasoning, like probabilistic approaches to the problem of confirmation or the assessment of probabilistic explanations. The issues associated with the lack of monotonicity of probabilities have been discussed under the label *the problem of the proper reference class*. Let us take a closer look at this problem.
The classical theory of probability assumes that probabilities can be assigned only to populations and not to single events (cf. Section 4.1.1). Thus, a statement like *the probability that this guy who is a have smoker lives more than 60 years is about 50%* means that this

guy is a member of the population where the median life expectation is 60 years. Consequently the chances to live longer than 60 years of a randomly drawn subject from this population are about 50%. This raises the question to which population or *reference class* a subject should be assigned to in order to infer information about her life expectancy. For example, if the guy in question is also a top athlete with a healthy diet who lives a life without stress, one might expect a higher life expectancy since this additional information indicates that he should be assigned to this specific sub-population that might have a higher life expectancy. Consequently, the reference class of heavy smokers may not be the correct one.
The problem of the proper reference class relates to conditional probabilities since the reference class (or population) is given by the conditioning set. To stay with our example, the conditional distribution:

$$P\left(\textit{years of life} \mid \textit{heavy smoker}\right),$$

might be quite different from the conditional distribution:

$$P\left(\textit{years of life} \mid \textit{heavy smoker}, \textit{top athlete}, \textit{healthy diet}, \textit{stressless life}\right).$$

Assigning a person to the wrong reference class amounts to considering an incorrect conditional probability. This can result in a wrong prediction and judgment, respectively. Here is a famous example.

Ex. 4-25: The case O. J. Simpson:
Alan M. Dershowitz, a Harvard professor and advisor of the O. J. Simpson defense team stated on U.S. television that only about 0.1% of wife batterers actually murder their wives and claimed that therefore evidence of abuse and battering should not be admissible in a murder trial.
The judgment of Dershowitz is based on the improper reference class. He uses as a reference class the population of *husbands who battened their wives*. According to his statement, the probability that a (randomly drawn) man from this population kills his wife is only 0.1% (or, equivalently, *0.1% of the men in this population kill their wives*):

$$P\left(\textit{killing his wife} \mid \textit{battened his wife}\right) = 0.001.$$

This ignores the fact that O. J. Simpson's wife has been killed. Thus the relevant reference class consists in the population of *husbands who battened their wives, and whose wives have actually been killed.* Thus the desired probability is:

$$P\left(\textit{killing his wife} \mid \textit{battened his wife}, \textit{his wife has been killed}\right).$$

Obviously, the second conditional probability is considerably higher than the one mentioned by Dershowitz since if the wife of men who battened her has been killed the probability is quite high that the murderer is the husband. On the other hand, the probability that a man who battens his wife actually kills her is low since most of these wives are not murdered.

Comment:

O. J. Simpson was accused to have killed his wife. However, he was acquitted by the jury despite the fact that the evidence that he had murdered his wife was overwhelming. Later on the members of the jury admitted to have committed an error.

The problem of the proper reference class can also be found in the context of regression modeling where it consists in the *problem of the specifying the correct model* or *the problem of model misspecification*. Regression modeling is concerned with the modeling of the conditional expectation $\mathrm{E}\left(Y \middle| \mathbf{X}=\mathbf{x}\right)$ where Y is the dependent variable and $\boldsymbol{X}$ is a vector of predictors (or independent) variables. The values of the predictor variables define the reference class of a single individual. Thus, the regression equation enables the computation of the expected value of the dependent variable for a person with the values $\mathbf{X}=\mathbf{x}$ on the predictor variables. With linear regression, the conditional expectation is models by means of a linear equation:

$$\mathrm{E}\left(Y \middle| \mathbf{X}=\mathbf{x}\right)=\beta_0+\beta_1\cdot x_1+\beta_2\cdot x_2+\cdots+\beta_k\cdot x_k$$

Usually, it is assumed that the values of the dependent variable for a specific population with values $\mathbf{X}=\mathbf{x}$ are normally distributed with mean $\mathrm{E}\left(Y \middle| \mathbf{X}=\mathbf{x}\right)$ and variance σ_Y^2 (identical for all populations). The latter can also be estimated on the basis of the regression model and its basic assumptions. Since the mean and the variance determine completely the (univariate) normal distribution the conditional distributeons of the dependent variable given a specific combination of the values of the predictor variables is completely specified.

Ex. 4-26: Conditional distributions in linear regression:

Given: A population where the relationship between two variables is described by the following linear regression model:

$$y=0.5\cdot x+\varepsilon \qquad \left[\varepsilon \sim N\left(0,25\right)\right]$$

Sample data were generated using the regression equation and the distribution of the error term ε. *Figure* 4-7 depicts the sampled data, the estimated regression line based on the sample data, and the conditional distributions of the dependent variable *y* given the values 0, 20, 40, 60, and 80 of the independent variable *x* (the red normal distribution curves). The means of the normal distributions are located on the regression line, and the conditional distributions all have the same standard deviation: $\sigma = 25$.

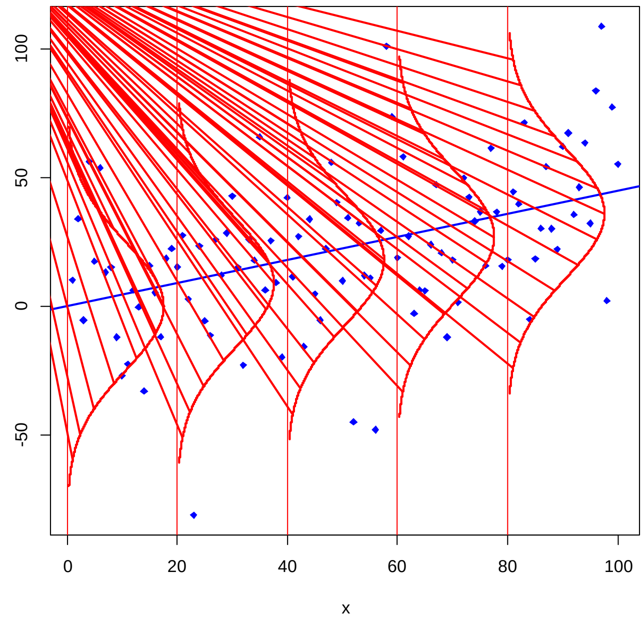


Figure 4-7: *Conditional distributions (represented by the red normal density curves) of the dependent variable y given values on the independent variable x.*

An important assumption underlying the model states that all relevant variables have to be included into the model. A variable is considered as relevant if it is statistically related to the other independent variables. However, the addition of a new variable to a set of existing ones defines new populations with each population given by a specific combination of values of the independent variables. Consequently, a specification error by not including relevant variables into the regression model can be interpreted as not considering the correct reference classes. The requirement of including all relevant variables into the regression model constitutes a special case of the following general principle:

Principle 4-3: *Taking the complete relevant information into account:*

Probabilistic judgments have to be based on the *complete relevant information* in the knowledge base.

A variable Z is irrelevant if it is conditionally independent of the variable Y if it is *conditionally independent* of the latter given the other variables $X_1, X_2, \ldots, X_k$ in the conditioning set:

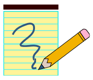
$$P\left(Y \middle| X_1, X_2, \ldots, X_k, Z\right) = P\left(Y \middle| X_1, X_2, \ldots, X_k\right)$$

Comment: Conditional independence of Y and Z given X means the Z does not change the conditional distribution of Y given X (cf. Section 4.3.1.4):

$$P\left(Y \middle| X = x, Z = z\right) = P\left(Y \middle| X = x\right) \qquad \left(\forall x \in X, \forall z \in Z\right)$$

In words:
Adding Z to the conditioning set does not change the distribution of Y; or:

The distribution of Y is identical in the sub-populations defined by the values of X and the sub-population defined by the combination of values of X and Y.



Notational Convention 4-1:

The symbol »$\forall$« is an abbreviation of »for all«. Thus the term $\forall x \in X$ stands for »for all values x of variable X« (and similarly for $\forall z \in Z$).

The problem of the proper reference class will be discussed further in the context of Bayesian reasoning (cf. Section 4.7.3.1).

4.3.1.7 The Influence of Causal and Diagnostic Reasoning on the Assessment of Conditional Probabilities

The assessment of conditional probabilities is influenced by causal considerations: People focus predominantly on the causal influence of an event on a future one thereby disregarding the diagnostic significance of a future event on the assessment of the probability of a previous event. A study of Tversky and Kahneman (1982b) demonstrates that this can lead to inconsistent probability judgments.

Ex. 4-27: Causal direction and the assessment of conditional probabilities (Tversky & Kahneman, 1982b, pp. 122-123):

Given: The following two problems:

Problem A: Which of the following probabilities is higher?

(i) The probability that, within the next five years, Congress will pass a law to curb mercury pollution, if the number of deaths attributed to mercury poisoning during the next five years exceeds 500.

(ii) The probability that, within the next five years, Congress will pass a law to curb mercury pollution, if the number of deaths attributed to mercury poisoning during the next five years *does not* exceed 500.

Problem B: Which of the following probabilities is higher?

(i) The probability that the number of death attributed to mercury poisoning during the next five years will exceed 500, if Congress passes a law within the next five years to curb mercury pollution.

(ii) The probability that the number of death attributed to mercury poisoning during the next five years will exceed 500, if Congress does not pass a law within the next five years to curb mercury pollution.

Result:

Most participants (140 of 166) estimated the probability of (i) in Problem A as higher than the probability of (ii), whereas for Problem B the order of estimated probabilities was reversed, i.e., the probability of (ii) was assessed as being greater than that of (i).

Inconsistent probability judgments:

The judgments of the majority of participants concerning the relative size of the probability for the two problems contradict the laws of probability. They are thus not consistent.

Formal analysis:

We use the following symbols:

C = Congress passes a law within the next five years to curb mercury pollution;

$\bar{C}$ = Congress does not pass a law within the next five years to curb mercury;

D = The number of death attributed to mercury poisoning exceeds 500;

$\bar{D}$ = The number of death attributed to mercury poisoning does not exceed 500.

The modal response pattern: (i) > (ii) for *Problem A* and (ii) > (i) for *Problem B* can thus be represented symbolically by the following pair of inequalities:

$$P\left(C|D\right) > P\left(C|\bar{D}\right) \text{ and } P\left(D|C\right) < P\left(D|\bar{C}\right).$$

This pattern of responses is in opposition to probability theory that demands the following equivalence (Exercise 4-10):

$$P\left(C|D\right) > P\left(C|\bar{D}\right) \Leftrightarrow P\left(D|C\right) > P\left(D|\bar{C}\right).$$

Interpretation:

The modal response pattern corresponds to the following line of reasoning:

(i) Many deaths attributed to mercury poisoning will induce Congress to pass a law against mercury pollution.

(ii) A law against mercury pollution will prevent many deaths due to mercury poisoning.

This line of reasoning is quite reasonable. However, it ignores the following diagnostic implication:

The presence of a law indicates fewer deaths and *the absence of a law indicates many deaths*.

By symmetry:

Fewer deaths indicate the presence of a law and *many deaths indicate the absence of a law.*

The modal response pattern indicates that participants' reliance on causal reasoning results in ignoring the diagnostic impact of one event with respect to the other.

Let us first illustrate why the inequality $P(C|D) > P(C|\overline{D})$ reflects the presence of causal reasoning:

1. Considering D as a cause of C implies that the probability $P(C|D)$ should be high since many deaths should lead to the passing of a law.
   If, on the other hand, D is regarded as a diagnostic indication of the presence of a law then $P(C|D)$ should be low since the presence of a law should prevent the appearance of many future deaths.
   Consequently, $P(C|D)$ is high if D is regarded as a cause of C whereas it is low if D is regarded as a diagnostic sign of C.
2. Analogously, $P(C|\overline{D})$ is low if D is regarded as a cause of C whereas it is low if $\overline{D}$ is regarded as a diagnostic sign of C.

   The fact that $P(C|D) > P(C|\overline{D})$ indicates that participants prefer the causal interpretation, paying less attention to the diagnostic aspect that the presence of a high and low death rate, respectively, provides an indication of whether there exists a law to curb mercury pollution.

A similar line of reasoning with C and $\bar{C}$ as the conditioning variable results in the inequality $P(D|C) < P(D|\bar{C})$ if participants prefer the causal interpretation:

1. In case of a causal interpretation $P(D|\bar{C})$ should be high since the absence of a law should cause many deaths.
   On the other hand, in case of a diagnostic interpretation, $P(D|\bar{C})$ should be low since the absence of a law indicates a low rate of deaths.
   Consequently, a causal interpretation results in a high value of $P(D|\bar{C})$ whereas a diagnostic interpretation should lead to a low value.
2. In case of a causal interpretation $P(D|C)$ should be low since the presence of a law should prevent many deaths.
   In case of a diagnostic conception $P(D|C)$ should be high since the presence of a law indicates a high death rate.
   Thus, a causal interpretation results in a low value of $P(D|C)$ whereas a diagnostic interpretation should lead to a high estimate.
   The fact that $P(D|C) < P(D|\bar{C})$ indicates that participants prefer the causal interpretation.

The reason for the inconsistency of the relations of conditional probabilities between Problem A and B arises because participants assign different causal roles to the two events for the two problems and the fact that they prefer the causal interpretation over the diagnostic one:

- *Problem A* induces people to treat a high or low death rate as the cause for the Congress passing or not passing a law, thus, $P(C|D) > P(C|\bar{D})$
- For *Problem B* the causal roles of the two events are reversed: The law passed (or not) by the Congress is conceptualized as the cause and the number of deaths is seen as the effect thus, $P(D|C) < P(D|\bar{C})$.

The example illustrates another important aspect: There exists a principal difference between human reasoning and reasoning based on probability calculus: For human beings causality as well as causal directions are important determinants for assessing the probability of events. Probabilistic computation based on the probability calculus is not concerned with causal directions. The only important aspect concerns the degree of stochastic dependence between variables (or the association between variables). It should however be noted that the causal structure provides cues with respect to the presence or absence of conditional independence between variables (cf. the discussion in Section 4.3.1.4).

4.3.2 Ignoring Base Rate Information

It has been noted in the context of the undervaluation of consensus information in Section 2.5.2.3 that consensus information is but a specific type of base rate information, and that its undervaluation is only a specific type of the more general phenomenon of base rate neglect.

Concept 4-4: *Base rate information*

Base rates represent the probability of occurrence of a specific event (or unit) within a reference class of events (a population).

By consequence, base rate represent the probability that a unit chosen randomly from a population is of a specific type (or contains a specific feature).

Examples:

- Proportion of women within the Swiss population [Reference class = Swiss population];
- Proportion of people with symptoms of depression among the Swiss students [Reference class = Swiss students].

Clearly, base rates contain relevant information about the presence or absence of a specific event. However, in a series of experiments Kahneman und Tversky (1973) demonstrated that base rate information is ignored or at least undervalued in the presence of additional information that is conceived of as being diagnostically relevant for the population in question.

Ex. 4-28: Base rate neglect (Kahneman & Tversky, 1973):

Given:

The following description:

Jack is a 45-year-old man. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies which include home carpentry, sailing, and mathematical puzzles.

On the basis of the description (plus base rate information, see below) participants had to provide the following estimate:

The probability that Jack is one of 30 (70) engineers in the sample of 100 is _______%.

For half of the participants (N = 85) 30 out of 100 persons were engineers and the other 70 were lawyers.

For the other half of participants (*N* = 86) the relative proportions of engineers and lawyers was reversed: 70 engineers vs. 30 lawyers.

Following the description (and estimation task) participants of both groups encountered the *null description*:

Suppose now that you are given no information whatsoever about an individual chosen at random from the sample.

The probability that the man is one of 30 engineers in the sample of 100 is _______%.

Results:

❑ In the presence of the full description of the person the manipulation of base rates (30/70 vs. 70/30) had a small (but significant) effect on the estimated probabilities only: The average estimates (that Dick was an engineer) in the (30/70) group was 50% and 55% in (70/30) group.

❑ For the null description the base rate is adequately taken into account.

❑ If the null description is replaced by the following completely uninformative description:

Dick is a 30-year-old man. He is married with no children. A man of high ability and motivation, he promises to be quite successful in his field. He is well liked by his colleagues.

the base rates are ignored completely: The median estimate was in both groups 50%.

The example demonstrates that people ignore base rate information in the presence of more specific information that may be of more or less diagnostic value.

There is however a situation where base rates have, in generally, an impact on the inference. In this case *causal base rates* are employed suggesting a causal factor that is associated with the different frequencies, and, in addition, the causal factor is also relevant for the estimation of the probability of the target event.

Ex. 4-29: Causal base rates (Ajzen, 1977, based on Tversky & Kahneman 1982b):

Given:

❑ A description of the academic achievements of a student.

❑ Two description of the situation:

(a) *Two years, ago a final exam war given in a course at Yale University. About 75% of the students failed (passed) the exam.*

(b) Two years ago, a final exam was given in a course at Yale University. An educational psychologist interested in scholastic achievements interviewed a large number of students who had taken the course. Since he was primarily concerned with reactions to success (failure), he selected mostly students who had passed (failed) the exam. Specifically, about 75% of the students in his sample had passed (failed) the exam.

The base rates in the two descriptions of the two situations differ with respect to their causal role:
The base rates in (a) implicate a causal factor that permits an explanation of the rate of success (failure) within the population: The difficulty of the exam. Obviously in case of a success rate of 75% the exam was easier than in the case of a rate of 25%.
In the description (b) no such causal factor is suggested since students are selected by the psychologist according to their rate of success.

The task of participants consisted in assessing the probability of success of a target person whose academic ability was shortly described.
Result:
As expected, the impact of the base rate information was different for the two descriptions:

- ❑ For the description in (a) the difference of the estimates between 75% vs. 25% success was 34%.
- ❑ For the description in (b) the difference of the estimates between 75% vs. 25% success was only 12%.

Interpretation:
In case of description (b) participants' judgments were based predominantly on the description of the academic ability of the target person, whereas in case of description (a) the base rate of success in the exam was more strongly taken into account.

In conclusion, base rates are, in general, ignored in the presence of more specific (pseudo-) diagnostic information except for the case of base rates exhibiting causal implications.

4.3.3 Risk Communication and Risk Assessment

One important practical problem that is related to the understanding of probability information concerns the public communication of risks. Specifically in the field of medicine and health the numbers communicated are not well understood by the public or they are of little use in conveying risks (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008).

Obviously the format of how probability information about possible risks is presented can make a great difference. For example, a news report stating:

> *33% of the Swiss people will be infected by the dangerous virus XYZ.*

will probably receive less resonance than the message:

> *1 out of every 3 Swiss people will be infected by the dangerous virus XYZ.*

A study of Slovic, Monahan, & MacGregor (2000) demonstrated the differential effect of the way in which probabilistic information had to be provided by experts.

Ex. 4-30: Frequentist vs. probabilistic information format (Slovic, Monahan & MacGregor, 2000).

Members of the American Academy of Psychiatry and Law received real cases of violent patients, summarized on a single page.

- ❑ The first group of participants had to assess the probability that Mr. Jones will commit an act of violence in case of being released on license.
- ❑ The second group had to assess how many people out of 100 people like Mr. Jones will commit an act of violence in case of being released on license.

The probability judgments were distinctly higher than the frequency judgments.

The subsequent example demonstrates a problem concerning risk communication in medicine.

Ex. 4-31: Risk communication in medicine (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008).

In October 1995, the U.K. Committee on Safety of Medicines issued a warning that third-generation oral contraceptive pills increased the risk of potentially life-threatening blood clots in the legs or lungs twofold – that is, by 100%.
The news caused great anxiety, and distressed women stopped taking the pill, which led to unwanted pregnancies and abortions.

How big is 100%? The studies on which the warning was based had shown that of every 7,000 women who took the earlier, second-generation oral contraceptive pills, about 1 had a thrombosis; this number increased to 2 among women who took third-generation pills. That is, the absolute risk increase was only 1 in 7,000, whereas the relative increase was indeed 100%.
Absolute risks are typically small numbers while the corresponding relative changes tend to look big – particularly when the base rate is low. Had the committee and the media reported the absolute risks, few women would have panicked and stopped taking the pill.
Ironically, abortions and pregnancies are associated with an increased risk of thrombosis that exceeds that of the third generation pill.

In epidemiology, two measures of risk are prevailing: The *relative risk* and *odds ratios* (Kahn & Sempos, 1989).

Concept 4-5: *Relative Risk and Odds Ratio*

Tab. 4-3 presents a contingency table with the entries containing the number of the joint occurrence of a risk factor and a disease. For example, Cell *A* contains the number of cases for which the disease and the risk factor are both present.

Tab. 4-3: *Contingency table containing information about the joint occurrence of a disease and a risk factor.*

	Disease		
Risk factor	Yes	No	Σ
Present	*A*	*B*	*A*+*B*
Absent	*C*	*D*	*C*+*D*
Σ	*A*+*C*	*B*+*D*	*N*

The relative risk (*RR*) is given by the formula:

$$RR = \frac{A/(A+B)}{C/(C+B)},$$

where the letters represent the number of cases in the different cells of the contingency table.
Thus, *RR* represents the probability of the disease being present given the risk factor is present divided by the probability of the disease in the population where the risk is not present:

$$RR = \frac{P(Disease\ present \mid Risk\ present)}{P(Disease\ present \mid Risk\ absent)}$$

The odds ratio is given by the equation:

$$OR = \frac{A \cdot D}{B \cdot C}.$$

The odds ratio represents the odds of the disease being present vs. absent for the sub-population being exposed to the risk, divided by the odds of the disease being present vs. absent for the sub-population without risk:

$$OR = \frac{A/B}{C/D}.$$

Alternatively, the odds ratio can be interpreted as representing the odds of the risk factor being present vs. absent for the sub-population with the disease, divided by the odds of the risk factor being present or absent for the healthy sub-population:

$$OR = \frac{A/C}{B/D}.$$

Both quantities, *RR* and *OR*, are measures of the strength of association between the two variables. Due to this characteristic they are used in epidemiology.

Comment:

OR is closely related to Yule's that has been considered as a measure of association between two binary variables in contingency tables *Q* (cf. Method 2-1 on page 29). In fact, *Q* is a function of *OR* (and vice versa):

$$Q = \frac{OR-1}{OR+1}, \qquad OR = \frac{1+Q}{1-Q}.$$

In case of large populations and rare diseases *RR* and *OR* are nearly identical. For various reasons, OR is the measure to be preferred (cf. Kahn & Sempos, Chapter 3).

Ex. 4-32: Relative risk (RR) and odds ratio (OR) in epidemiology (Kahn & Sempos, 1989).

Tab. 4-4 depicts the joint frequencies of smoking and stroke form male participants.

Tab. 4-4: *Twelve-Year risk of stroke among male smokers and non-smokers (from Kahn & Sempos, 1989, page 46).*

	Stroke		
Smoker	Yes	No	Σ
Yes	171	3264	3435
No	117	4320	4437
Σ	288	7584	7872

Using these data we get:

$$RR = \frac{171/3435}{117/4437} = 1.89 \text{ and } OR = \frac{171 \cdot 4320}{117 \cdot 3264} = 1.93.$$

Thus, as one would expect, there is a positive relationship between smoking and strokes, with the probability of getting a stroke being nearly double the size of the repective probability for non-smokers.

RR and *QR*, are certainly valuable as measures of the strength of association between rsik factors and diseases. However, in case of rare diseases as in Ex. 4-31 they can be quite misleading. For this example both measures assume a value of about 2 indicating a clear association between third generation pills and thrombosis. However, since the absolute risk is only about $2/7000$ for women taking the pill, the positive association is of little relevance for practical purposes.

In conclusion, in order to properly assess the risk of a possible risk factor it is useful to get information not only of the relative risk but to also take the absolute risk into consideration.

Gigerenzer et al. (2008) also provide a nice example of the improper application of an important measure used in epidemiology, the *survival rate*:

> In a 2007 campaign advertisement, former New York mayor Rudy Giuliani said, "I had prostate cancer, 5, 6 years ago. My chance of surviving prostate cancer – and thank God, I was cured of it – in the United States? Eighty-two percent. My chance of surviving prostate cancer in England? Only 44 percent under socialized medicine" […]. For Giuliani, these health statistics meant that he was lucky to be living in New York and not inYork, since his changes of surviving prostate cancer appeared to be twice as high. This was big news. As we will explain, it was also a big mistake. (Gigerenzer et al., 2008, p.53).

The problem of using the survival rate for comparing the mortality in different countries lies in the fact that survaivel rate depends on two factors:

(a) The time of a diagnosis: Younger patients with the disease that have been diagnosed earlier usually live longer.

(b) The specificity (or false alarm rate) of the diagnostic instrument.

The survival time *ST* for a specific survival interval, say 5 years, is defined as follows:

$$ST_{5\text{ years}} = \frac{\#\text{ Positive diagnosis and still alive 5 years after the diagnosis}}{\#\text{ Positive diagnosis}},$$

where the symbol # indicates *number of people*.

Thus, if in the diagnostic systems for prostate cancer differ between the United States and Great Britain a comparison of survival rates is not sensible. In fact, according to Gigerenzer et al. (2008), the systems differ: In Great Britain the diagnosis is based on symptoms whereas in

the United States prostate-specific antigen (PSA) blood test is used. This results in earlier diagnosis of prostate cancer as well as in higher false rates. Both factors cause higher survival rates for the United States.
The inutility of prostate cancer screening using PSA, digital rectal examination (DRE), and transrectal ultrasound (TRUS) guided biopsy is exhibited by a meta analysis of the Cochran collaboration. Tab. 4-5 shows the results. Obviously, screening has no effect on the mortality rate due to prostate cancer or any other disease (cf. first and second row of the table. However, with screening significant more men diagnosed as having prostate cancer. It can be assumed that most of these additional cases constitute false alarms.

Tab. 4-5: *Results of prostate cancer screening for men of 45 to 80 years (Ilic, Neuberger, Djulbegovic & Dahm, 2013)*

	Screening			
Outcome	No	Yes	*Odds* [95% CI]	*N*
All-cause mortality	21 per 100	21 per 100	1.0 [0.96-1.03]	294,856
Prostate cancer-specific mortality	7 per 1000	7 per 1000	1.0 [0.86-1.17]	341,342
Prostate cancer diagnosis	68 per 1000	88 per 1000	1.3 [1.02-1.65]	294,856

Notes:
Odds = odds of proportions between screening and not screening
95% CI = 95% confidence interval,
N = number of participants.

The situation constitutes a typical case of confounding. In addition, it is another illustration of the problems involved in causal reasoning (cf. Section 2.6.2): The confounding factor is the *type of diagnosis* that is correlated with the factor *country* and exerts an influence on the dependent measure, the *survival rate*. Due to this confounding, longer survival rates cannot be attributed to a better health system of the United States as indicated by Giuliani.
Gigerenzer et al. (2008) propose to compare the mortality rates (*MR*) of the two countries instead of survival rates. The mortality rate for a specific time interval, say 1 year is given by the expression:

$$MR_{1\text{ year}} = \frac{\text{Number of death in the group due to cancer in 1 year period}}{\text{Number of people in the group}}$$

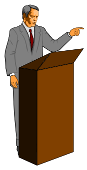
Gigerenzer et al. (2008) report a mortality rate due to prostate cancer of 26 and 27 per 100,000 for the United States and Great Britain, respectively.
Unfortunately, the mortality rate has a similar problem as the survival rate, the presence of a possible confounding variable. In this case the

variable *age* might be a confounder. If, for example, the age distribution is different for the two countries, and, consequently, *age* is correlated with *country* then the crude mortality rate cannot be used for comparsion. Instead, mortality rates adjusted for differences in the age distribution have to be employed (cf. Kahn & Sempos, 1989, Chapter 5). However, as exhibited by the data in Tab. 4-6 the age structure of the two countries is similar. Consequently, age seems not to be a confounder.

Tab. 4-6: *Estimated age distribution of the year 2017 for the United Kindom and the United States.*

	Country	
Age category in years	United Kindom	United States
0-14	17.53%	18.73%
15-24	11.90%	13.27%
25-54	40.55%	39.45%
55-64	11.98%	12.91%
65+	18.04%	15.63%

Source:
https://www.indexmundi.com/factbook/compare/united-kingdom.united-states



Comment 4-5:

It might be argued that the age distribution of the whole population is not relevant here. Of greater importance is the age distribution of the sub-population with prostate cancer.

This argument is correct. However, the actual presentation just wanted to show the importance of taking age into account in the comparison of mortality rates from different countries.

Another type of erroneous risk communication in medicine consists in the confusion of the sensitivity of the diagnosis, i.e., the probability of a positive diagnosis given the presence of the disease, with the inverse conditional probability, i.e., the probability of a disease given a positive diagnosis (cf. Section 4.3.1.5). This results in unjustified panic as well as people undergoing unnecessary and possibly harmful treatments.

4.4 Probabilistic Reasoning

Biases and problems in probabilistic reasoning have been investigated predominantly in the context of Bayesian reasoning where the judgments of humans are compared to results of applying Bayes theorem that is considered as the normative standard.

In the following, we present one famous Bayesian problem: The cab problem of Tversky and Kahneman. Using this problem, different representations of Bayesian problems and the Bayes theorem, respectively, will be illustrated:

(a) The tree format,
(b) The sampling representation,
(c) The contingency table format,
(d) Bayes formula, and
(e) Bayes theorem in odds format.

This sets the stage for for the next step: The illustration of the structure of probabilistic reasung in general. This enables us to conceptualize probabilistic reasoning as a problem solving activity that fits the problem space approach to problem solving. In addition, it will be shown that Bayesian reasoning is but a special case of probabilistic reasoning that consists in the application of a specific sequence of probabilistic operations.
We then present different psychological explanations of peoples' behavior in probabilistic reasoning tasks. Specifically the controversy between evolutionary psychologists and problem theorists will be discussed. In addition, fallacious probabilistic intuitions and heuristics resulting in errors of probabilistic reasoning are described and illustrated using the famous Monty Hall problem.

4.4.1 Bayesian Problems and Bayes Theorem

4.4.1.1 The Cab Problem

A classical problem demonstrating that peoples' probabilistic reasoning is not in agreement with Bayes theorem is the so called *cab problem* of Tversky and Kahneman (1982c).

Ex. 4-33: The Cab Problem (Tversky & Kahneman 1982c):

Description of the problem:

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85% of the cabs in the city are Green (G) and 15% are Blue (B).
(b) A witness identified the cab as Blue (»B«). The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

What is the probability that the cab involved in the accident was Blue rather than Green?

Result: The median and modal answer was 0.80.

Interpretations:

1. Participants' answers indicate a confusion of the following two conditional probabilities:

   $$P\left(Cab\ is\ blue \middle| Witness\ identifies\ cab\ as\ blue\right) = 0.41,$$

   and

   $$P\left(Witness\ identifies\ cab\ as\ blue \middle| Cab\ is\ blue\right) = 0.80.$$

   The asked for target probabilitiy is the first one of the two conditional probabilities.

2. The result can also be interpretated as an indication of base rate neglect: The modal answer ignores the relative frequencies of the cabs in town.

In order to solve the problem one could use Bayes formula. However, this provides no insight of why the looked for probability is 0.41 and not 0.80. There exist various attempts to teach Bayesian reasiong using different representations of the problem (see, e.g., Böcherer-Lindner & Eichler, 2019; Hoffrage, Krauss, Martignon, & Gigerenzer, 2015; Sedlmeier & Gigerenzer, 2001; Talboy & Schneider, 2017). In the following five different representations of the cab problem are discussed: Outcome trees, a sampling representation, contingency tables, Bayes formula, and Bayes theorem in odds format.

4.4.1.2 Representation of the Cab Problem by Means of Outcome Trees

Outcome trees provide a convenient way to represent and analyze simple Bayesian problems like the cab problem. The outcome tree can be applied with frequencies or probabilities. Various studies demonstrate that using frequencies improves performance as well as memory for solutions (see, for example, Hoffrage, et al. 2015; Sedlmeier & Gigerenzer, 2001). Thus, we first present the outcome tree in frequency format.

Ex. 4-34: The cab problem represented by means of an outcome tree in frequency format:

Given: The following symbols for denoting different events:

B = The cab is blue.

G = The cab is green.

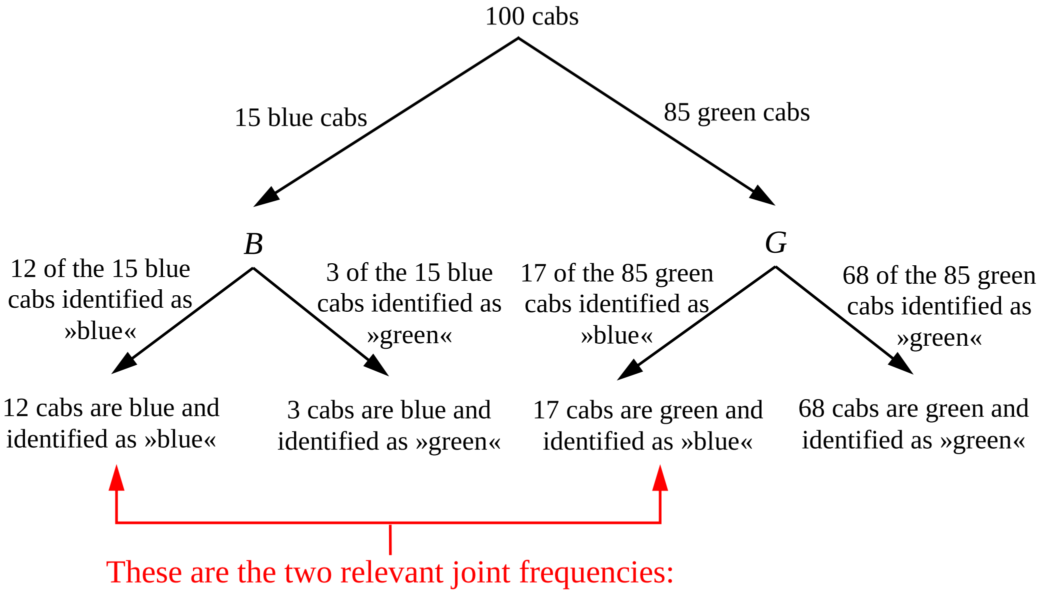
»*B*« = The witness identified the cab as blue.

»*G*« = The witness identified the cab as green.

Figure 4-8 depicts the outcome tree representation of the problem in frequency format.

The tree partitions the set of 100 cabs in two steps:

1. The whole set of 100 cabs is partitioned into the number of blue and green taxis.
2. The second step partitions each of these two sets in two subsets consisting of the number of cabs that have been identified as »blue« and »green« by the witness.



$$P\left(B \mid »B«\right) = \frac{\#\left(B, »B«\right)}{\#\left(B, »B«\right) + \#\left(G, »B«\right)} = \frac{12}{12+17}$$

Figure 4-8: *Representation of the cab problem by means of an outcome tree in frequency format.*

The stepwise partitioning of the set of 100 cabs results in the joint events and their frequencies. The Venn diagram in Figure 4-9 illustrates the partitioning of the sampling space (population) for the cab problem.

The looked for probability is the conditional probability $P\left(B \mid »B«\right)$ that the cab is blue given that the witness had identified the cab as »blue«. Thus, the final step consists in applying the definition of conditional probability in terms of joint and marginal probabilities (cf. Concept 4-1 on page 123 and Ex. 4-13 on page 124):

$$P\left(B \mid »B«\right) = \frac{P\left(B \wedge »B«\right)}{P\left(»B«\right)} = \frac{\#\left(B \wedge »B«\right)/\#N}{\#\left(»B«\right)/\#N}$$

$$= \frac{\#\left(B \wedge »B«\right)}{\#\left(»B«\right)} = \frac{\#\left(B \wedge »B«\right)}{\#\left(B \wedge »B«\right) + \#\left(\overline{B} \wedge »B«\right)}$$

$$= \underline{\underline{\frac{12}{12+17}}}$$

In the equation the symbol # denotes the number of case and *N* denotes the size of the whole population (in our case *N* = 100). Note that the size *N* cancels in the computation of the conditional probabilities. Consequently, one arrives at the same results with any size (e.g. *N* = 1000). The tree representation in frequency format is closely related to the natural sampling representation of the problem (cf. Section 4.4.1.3).
The problem can also be represented by means of an outcome tree that uses probabilities instead of frequencies.

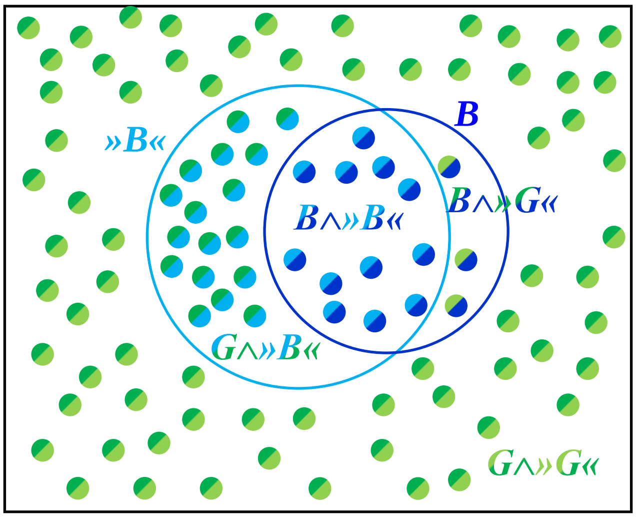


Figure 4-9: *Venn diagram illustrating the partitioning of the sample space for the cab problem (B = blue cab, »B« = witness identifies the cab as »blue«).*

Ex. 4-35: The cab problem represented by means of an outcome tree in probabilitiy format:

Given: The following symbols for denoting different events:

Ω = The whole population of cabs
B = The cab is blue.
G = The cab is green.
»*B*« = The witness identified the cab as »blue«.
»*G*« = The witness identified the cab as »green«.

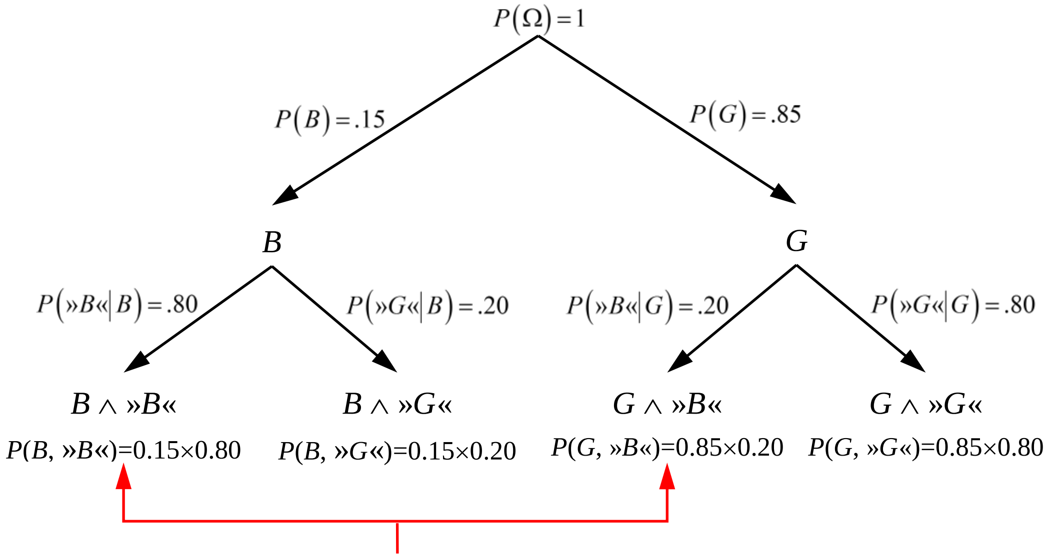
Figure 4-10 depicts the outcome tree representation of the problem.

The outcome tree with probabilities looks similar to the on in frequency format in Figure 4-8. There are two differences:

1. Instead of the number of cases at root of the frequency tree the probability contains the number 1. By consequence the sizes of the various sets are measured (or represented) by means of probabilities and not frequencies. As noted above, the whole number of cases cancels in the computation of the conditional probabilities. It thus does not matter which number is used.

2. A more important difference concerns the the probabilities and frequencies, respectively that are associated with the arrows pointing to the leaves of the trees. In case of the probability trees these represent conditional probabilities, whereas in case of the frequency tree these quantities are called *natural frequencies*.

The important difference consists in the fact, that the conditional probabilities do not contain information concerning the size of the sub-population that is involved. This is due to the fact that conditional probabilities represent the proportion of case with the target event within the sub-population. This proportion is independent of whether the sub-population comprises many members or only few. For this reasons, the conditional probabilities are *normalized* quantities.



These are the two relevant joint probabilities:

$$P(B|»B«)=\frac{P(B,»B«)}{P(B,»B«)+P(G,»B«)}$$
$$=\frac{0.12}{0.12+0.17}$$

Figure 4-10: *Representation of the cab problem by means of an outcome tree in probability format.*

By contrast, *natural frequencies* incorporate the information about the size of the sub-population. For example, in Figure 4-8 the label *12 of 15 blue cabs identified as »blue«* contains the information about the number of blue cabs, i.e., the information about the size of the sub-population. Natural frequencies are thus not normalized (cf. Section 4.4.3, for different explanations why the representation of the cab problem in terms of natural frequencies simplifies the process of Bayesian reasoning).

It is also possible to use outcome trees with normalized frequencies, e.g. *80 of 100 blue cabs are identified as »blue«*. In this case the statement provides the proportion of cases with respect to the size of the whole population. Information about the size or the sub-population is no longer available.

Note also that the indicated frequency 12 in the natural frequency statement refers to the joint frequency, i.e. to the number of taxis in the population that are blue and identified as »blue«. This number represents the *number of hits* for the blue cabs and not the *hit rate* (0.80), as is the case with the probability tree. Similarly, the number 17 represent the *number of false alarms* for the green cabs and not the *false alarm rate* (0.20) that is shown with the probability tree. The number of false alarms corresponds to the joint frequency of green cabs in the population that are identified as blue.

Due to the fact that the probability tree presents the conditional probabilities only, the joint probabilities (at the leaves of the tree) have to be computed using the definition of conditional probabilities, e.g.:

$$P\left(B \middle| »B«\right) = \frac{P\left(B \wedge »B«\right)}{P\left(»B«\right)} \Rightarrow P\left(B \wedge »B«\right) = P\left(B \middle| »B«\right) \cdot P\left(B\right).$$

In the present case: $0.12 = 0.80 \cdot 0.15$.

With natural frequencies this is not required since, as already noted the first number mentioned in the natural frequency statement corresponds to the joint frequency. Had we used non-normalized frequencies (*80 out of 100*) the joint frequencies were not available and would thus have to be computed. Note that this is more tedious than in case of probabilities, specifically if a references size other than 100 was used as denominator, e.g., *1080 out of 1350*.

The tree representations reveal why the conditional probability that the cab is blue given that it has been classified as blue is only 41%: Due to the different base rates the number of erroneously green taxis classified erroneously as blue (*the number of false alarms*) is greater than the number of blue taxis classified correctly as blue (*the number of hits*).

4.4.1.3 Representation of the Cab Problem from a Sampling Perspective

Another illuminating representation of the problem provides a sampling perspective. The whole sampling process may be conceived of as consisting of three separate stages (Figure 4-11):

1. In the first stage a sample of 100 balls is drawn *with replacement* from an urn containing blue and green balls representing blue and green cabs. The relative number of balls in the urn reflects the relative number of cabs in town. In Figure 4-11 the urn contains 3 blue and 17 green balls (Alternatively, the urn might contain 15 blue and 85 green balls). The resulting sample of 100 balls contains the test cases that are used to assess the reliability of the witness. It contains (approximately) 15 blue and 85 green balls.

2. In the second stage, the process of classification is modeled using the balls sampled in the first stage. A ball is chosen from the sample and classified as »blue« (or »green«) according to the given proportions of correct and incorrect classifications: 0.8 of the blue and 0.2 of the green cabs are classified as »blue«. The respective balls are put into the respective urn (in the middle of Figure 4-11), according to the given proportions of correct and incorrect responses. If all balls have been classified, this urn contains approximately 12 blue and 17 green balls.

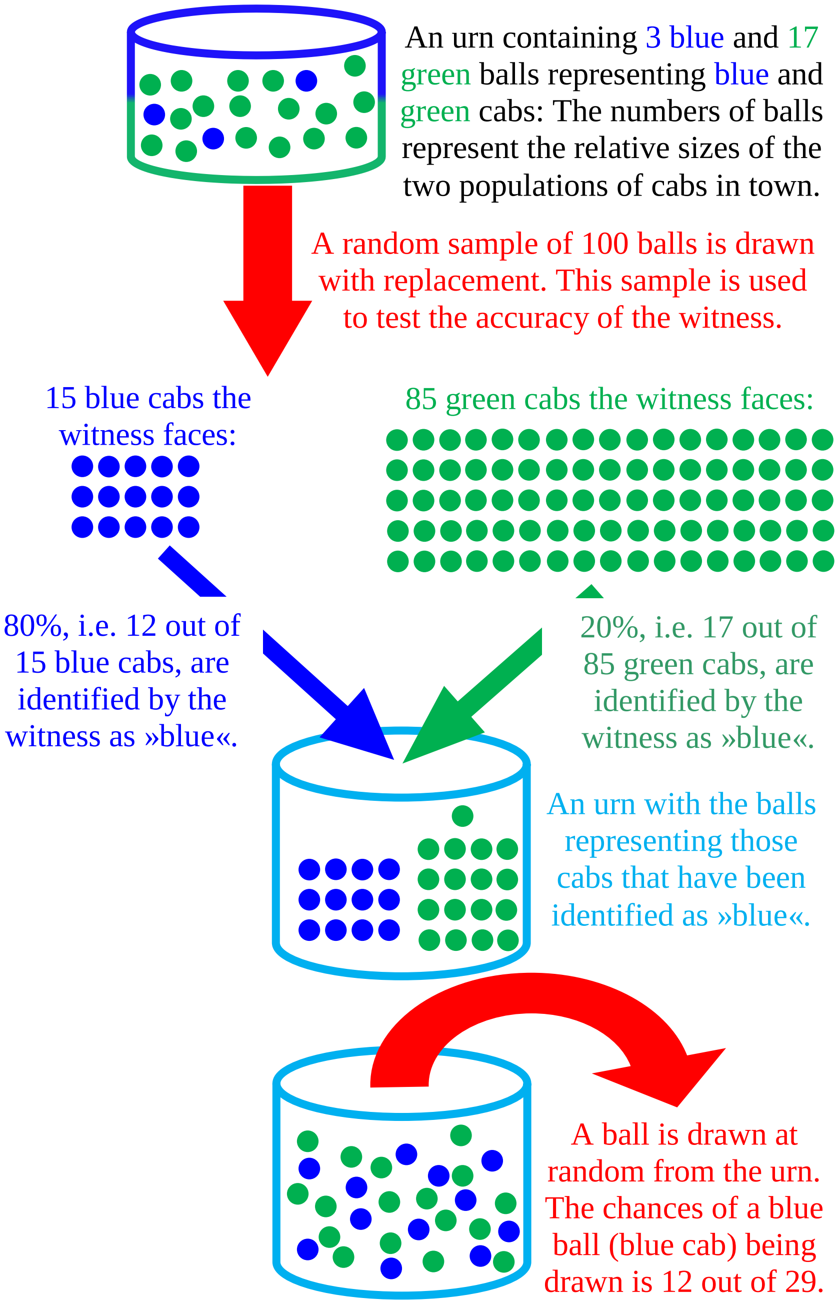


Figure 4-11: *Representation of the cab problem from a natural sampling perspective.*

3. In the final step a ball is drawn at random from the urn with the balls representing the cabs classified as »blue«. The chances that the ball drawn is blue is: $12/29 = .41$.

Note that in reality the color of the balls within the urn representing the cabs identified as »blue« cannot be distinguished. Thus it cannot be decided whether the judgment is a *hit* (i.e. the cab is really blue cab) or a *false alarm* (i.e. the cab is in fact green). One can only determine the probabilities of the two possibilities.

Similar to the outcome tree the sampling representation reveals why the probability of the cab being blue is only about 41% despite the relatively high accuracy of the witness (80%): Since there are many more green taxis than blue ones that have to be classified, more green than blue balls are put into the urn from which a final random draw is performed.

It may be argued that the described sampling process is unduly complex. The whole procedure may be conceptualized in a much simpler way:

1. Draw a ball from the first urn containing the balls that represent the blue and green cabs, with the number of balls representing the relative sizes of the cabs in town.
2. Perform the identification of the drawn ball according to the given probabilities.
3. If the the cab is identified as »blue« then record the color of the ball.

If this sampling process is repeated many times, (about) 41% of the balls whose color is registered will be blue and the other 59% will be green.

This latter version of the sampling has the disadvantage that, contrary to the original version, the involved populations are not directly discernible. In the version illustrated in Figure 4-11 the first urn represents the population of cabs in town, and the second urn represents the sub-population corresponding to the relevant conditioning event, i.e., those cabs that have been identified as »blue«. On the basis of the urn's composition the relevant conditional probability can be recovered by deviding the number of blue balls by the whole number of balls in the urn. The first urn represents the *prior odds*, that is, the proportion of blue to green cabs (3/17) previously to the process of Bayesian updating. The second urn represents the *posterior odds*, i.e., the proportion of blue and green cabs after the process of updating (12/17).

Both versions of the sampling process illustrate, however, the process of conditioning: In the first version, the conditioning consists in putting only those balls into the second urn that have been identified as »blue« thus constructing the relevant sub-population. In the second version the process of conditioning is realized by recording the color of those balls only that have been identified as »blue«, ignoring the other ones.

Let us now look at another representation of the problem.

4.4.1.4 Representation of the Cab Problem by Means of Contingency Tables

Another usefult representation of Bayesian problems concerns the presentation of frequencies and probabilities, respectively, by means of contingency tables. The contingency table contains the joint frequencies or probabilities of the events involved. Tab. 4-7 depicts the joint and marginal frequencies for the cab problem, whereas Tab. 4-8 exhibits the associated joint and marginal probabilities.

The tables may be constructed step-by-step from the description of the problem:

1. In the first step, the *marginal* frequencies/probabilities of blue and green cabs may be inserted in the left-most column of the table.
2. In the second step, these marginal probabilities are »split« into the joint frequencies/probabilities according to the conditional probabilities (similar to the splitting performed in the context of the outcome trees).

Tab. 4-7: *Contingency table representing joint and marginal frequencies.*

	Classification of witness		
Color of the cab	»blue«	»green«	Σ
blue	12	3	15
green	17	68	85
Σ	29	71	100

Tab. 4-8: *Contingency table representing joint and marginal probabilities.*

	Classification of witness		
Color of the cab	»blue«	»green«	Σ
blue	0.12	0.03	0.15
green	0.17	0.68	0.85
Σ	0.29	0.71	1.00

For example, using the frequency contingency table, 80% of the 15 blue cabs are identified as »blue«, and 20% are identified as »green«. Consequently, the 15 cabs have to be split according to the proportion of 4:1, resulting in 12 blue cabs identified as »blue« and 3 blue cabs indentified as »green«. Similarly, for the green cabs the respective pro-

portion (of »blue« vs. »green«) is 1:4, resulting in 17 green cabs identified as »blue« and 68 green cabs indentified as »green«.
The »splitting« of the marginal probabilities according to the proportions given by the conditional probabilities is less obvious than in case of outcome trees. Thus, the computation of the joint frequencies/probabilities seems to be easier with outcome trees (note that the joint frequencies/probabilities are located at the leaves of the trees).
By contrast, the computation of the required conditional probability seems to be more straightforward with contingency tables: Summing the rows of the table provides the (marginal) frequencies/probabilities of the conditioning events. These are shown in the last row of Tab. 4-7 and Tab. 4-8, respectively. Devision of the conditional frequencies/-probabilities by these marginal probabilities results in the table of conditional probabilities $P(C|W)$ of the color of the cabs given the testimony of the witness (Tab. 4-9). Note that the entries in the rows of the table must sum to one. The probability in question ($12/29$) is found in the upper left entry of the table. However, the table also provides the conditional probability of the cab being blue if the witness had identified the cab as »green«: $3/71$, as well as the other conditional probabilities of the color of the cab, given the testimony of the witness.

Tab. 4-9: *Table of conditional probabilities of the color of the cab given the testimony of the witness.*

	Classification of witness	
Color of the cab	»blue«	»green«
blue	$12/29$	$3/71$
green	$17/29$	$68/71$
Σ	1	1

Thus it might be concluded that outcome trees seem to be more favorable for computing joint frequencies/probabilities whereas contingency tables are more convenient in computing the relevant conditional probabilities on the basis of the joint probabilities or frequencies. Let us now turn to algebraic representations of the problem.

4.4.1.5 Bayes Formula

The representation of the cab problem and its solution by applying the Bayesian formula rests on the definition of conditional probability (cf. Section 4.3.1). In the present case, using the notation of the previous sections, the formula may be written as (cf. Section 4.4.1.2):

$$P(B|»B«) = \frac{P(B \wedge »B«)}{P(»B«)}.$$

The problem consists in the fact that the quantities on the right-hand side of the equation are not given, directly. However, as already exhibited in Section 4.4.1.2, both quantities can be computed from the given problem specification. Specifically, applying the definition of the conditional probability, we get:

$$P(B \wedge \text{»}B\text{«}) = P(B \mid \text{»}B\text{«}) \cdot P(\text{»}B\text{«}),\ \text{and}$$

$$P(G \wedge \text{»}B\text{«}) = P(G \mid \text{»}B\text{«}) \cdot P(\text{»}B\text{«})$$

Moreover, the probability $P(\text{»}B\text{«})$ is given by summing the two joint relevant probabilities (cf. the Venn diagram in Figure 4-9 (page 162) and the subsequent discussion, as well as the contingency table representation of the problem):

$$P(\text{»}B\text{«}) = P(B \wedge \text{»}B\text{«}) + P(G \wedge \text{»}B\text{«}).$$

We are now able to put the single pieces together:

$$\begin{aligned} P(B \mid \text{»}B\text{«}) &= \frac{P(B \wedge \text{»}B\text{«})}{P(\text{»}B\text{«})} = \frac{P(B \wedge \text{»}B\text{«})}{P(B \wedge \text{»}B\text{«}) + P(G \wedge \text{»}B\text{«})} \\ &= \underline{\underline{\frac{P(B \mid \text{»}B\text{«}) \cdot P(\text{»}B\text{«})}{P(B \mid \text{»}B\text{«}) \cdot P(\text{»}B\text{«}) + P(G \mid \text{»}B\text{«}) \cdot P(\text{»}B\text{«})}}} \end{aligned}.$$

The formula in the last row represents the standard version of Bayes theorem for the cab problem. This formula integrates the different steps that are involved in the construction the outcome trees and contingency tables, respectively, into a single algebraic expression. Thus, in order to solve the cab problem one has to identify the relevant probabilities in the problem description and insert it into Bayes formula.

However, unlike the outcome tree representations the application of Bayes formula does not provide an insight of why the resulting conditional probability $P(B \mid \text{»}B\text{«}) = 0.41$ is considerably lower then the inverse probability $P(\text{»}B\text{«} \mid B) = 0.80$.

In the context of Bayesian reasoning, a specific jargon is employed:

Notation 4-4: Bayesian terminology

1. The initial probabilities that are revised by applying Bayes theorem are called *prior probabilities* or *priors*. For the cab problem the prior probabilities are $P(B)$ and $P(G)$ that represent the distribution of cabs in town.
2. The conditional probabilities that are used for revising the prior probabilities are called likelihoods. For the cab problem the relvant likelihoods are: $P(\text{»}B\text{«} \mid B)$ and $P(\text{»}B\text{«} \mid G)$ i.e. the probabilities of the witness identifiying the cabs of different colors as »blue«.

3. The final (revised) probabilities are called *posterior probabilities*. They represent the probabilities after conditioning on the relevant event. In case of the cab problem the relevant posterior probability is $P(B|»B«)$, the probability that a cab identified as »blue« by the witness is really blue.

The notation employed in the context of Bayes theorem is due to fact that Bayes theorem is considered as a means for updating ones knowledge base due to new information (cf. Section 4.3.1.2, Ex. 4-2 on page 108, as well as Ex. 4-15 on page 130). The updating process replaces the prior by the posterior probabilities or, more generally, the prior by the posterior distribution. The latter may, in turn, constitute a prior distribution for a subsequent process of updating. With respect to the cab problem the process of updating is not obvious.
Regarding Bayes theorem as a device for updating probabilities, the latter are ususally associated with different hypotheses. Thus the prior probabilities are concerned with the probabilities of a set of hypotheses that are exhaustive and exclusive (cf. Figure 4-3 on page 129). The conditioning event *E* represents empirical evidence, and the posterior probabilities represent the probabilities of the hypotheses given the evidence *E*. In the case of the cab problem there are two hypotheses:

H_1: The cab is blue, and
H_2: The cab is green.

The conditioning event is the testimony of the witness that the cab was »blue«. The following specification describes Bayes theorem with *n* hypotheses:

Concept 4-6: *Bayes theorem:*

Given:

- ❑ A partition of the hypothesis space into a set of exhaustive and exclusive hypotheses: $H_1, H_2, \ldots, H_n$. This means that exactly one of the hypotheses is correct (cf. Figure 4-3 on page 129).
- ❑ The *priori* probabilities $P(H_1), P(H_2), \ldots, P(H_n)$ of the single hypotheses that sum to 1.0 (due to the fact that hypotheses are exhaustive and exclusive).
- ❑ An event *E* (an observation or some other form of evidence) that assumes on a specific probability under each of the given hypotheses: $P(E|H_1), P(E|H_2), \ldots, P(E|H_n)$. These conditional probabilities $P(E|H_i)$ of the event given the various hypotheses are called the *likelihoods*.

Comment: The likelihoods are not necessarily probabilities. They can also be densities. However, in the present context likelihoods are always probabilities.

Objective/result of the reasoning/computation:
The *posterior probabilities* of the hypotheses given the evidence *E*: $P(H_1|E), P(H_2|E), \ldots, P(H_n|E)$. These conditional probabilities represent the updated probabilities of the various hypotheses in the presence of evidence *E*.

Bayes theorem tells us that the required posterior probabilities $P(H_i|E)$ of hypotheses H_i given evidence *E* can be calculated according to the following equation:

$$P(H_i|E) = \frac{P(H_i, E)}{P(E)} = \underline{\underline{\frac{P(E|H_i) \cdot P(H_i)}{\sum_{j=1}^{n} P(E|H_j) \cdot P(H_j)}}}$$

Comments:

1. The Bayes formula provided above in the context of the cab problem is but a special case of the general one.
2. Bayes theorem is a normative rule that results directly from the axioms of probability theory and the definition of conditional probability. However, within the frequentist interpretation the prior probabilities $P(H_i)$ do not make sense in most contexts since the probabilities refer to single events (that are either true or false). Consequently, the application of the theorem in these contexts is not sensible.
3. In the philosophy of science, Bayesian confirmation theory has been established as the most influental theory of confirmation of scientific theories (see e.g. Hawthorne, 2011; Howson & Urbach, 2006).

The following example is intended to illustrate the principle aspects of probabilistic reasoning using Bayes theorem with three hypotheses.

Ex. 4-36: Bayesian reasoning / updating:
Given:
Three hypotheses concerning the bias of a coin:

$$H_1 : \pi_1 = 1/4$$

$$H_2 : \pi_2 = 1/2$$

$$H_3 : \pi_3 = 3/4$$

Where π_i denotes the (conditional) probability of the coin landing *Heads* under hypothesis H_i $(i=1,2,3)$:

$$P\left(Heads|H_i\right)=\pi_i \text{ and } P\left(Tails|H_i\right)=1-\pi_i.$$

It is assumed that the three hypotheses represent the only possibilities, i.e. they exhaust the space of hypotheses.
In addition, assume that the prior probabilities for each of the three hypotheses are the same, i.e.:

$$P(H_1)=P(H_2)=P(H_3)=1/3.$$

Now, the coin is tossed and lands *Heads* (*H*). What is the probability of the three hypotheses given the observed outcome *E*: $P\left(H_i|E\right)$, (i = 1, 2, 3).

Figure 4-12 depicts the sample space that consists of 6 sample points. Each point represents a combination of a hypothesis and an outcome. Note that the hypotheses are disjoint since each one covers different points in the space. In addition, the hypotheses are exhaustive, i.e. together they cover all points in the space.
The conditioning event is represented by the red ellipses. It covers the three points: $H_1 \wedge H$, $H_2 \wedge H$, and $H_3 \wedge H$. These points represent the combination with one of the three hypotheses being true and the actual outcome being *Heads* (*H*).
In order to compute the searched for probabilities: $P(H_i|E)$ we use Bayes theorem, specifically:

$$P(H_i|E)=\frac{P(H_i \wedge E)}{P(E)}$$

where (cf. *Figure* 4-12),

$$P(E)=P(H_1 \wedge E)+P(H_2 \wedge E)+P(H_3 \wedge E).$$

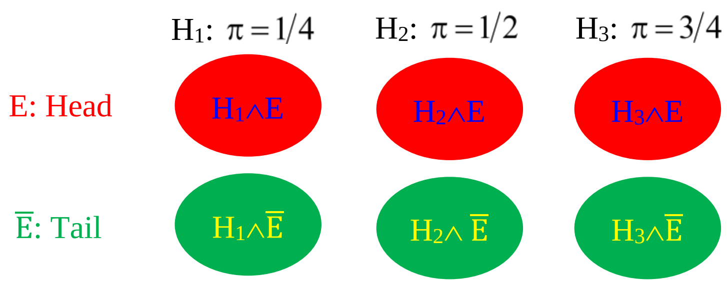


Figure 4-12: *Illustration of the sample space underlying the problem of Ex. 4-36: It consists of six joint events that result from combinating the three hypotheses H_1, H_2, and H_3 with the two possible outcomes: Heads and Tails.*

In order to calculate these quantities we first have to compute the joint probabilities $P(H_i \wedge E)$ (i = 1, 2, 3). This is easily accomplished using the equation:

$$P(H_i \wedge E) = \underbrace{P(E|H_i)}_{\pi_i} \cdot P(H_i) \quad (i = 1, 2, 3)$$

Thus,

$$P(H_1 \wedge E) = \underbrace{P(E|H_1)}_{\pi_1} \cdot P(H_1) = \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{12},$$

$$P(H_2 \wedge E) = \underbrace{P(E|H_2)}_{\pi_2} \cdot P(H_2) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6},$$

$$P(H_3 \wedge E) = \underbrace{P(E|H_3)}_{\pi_3} \cdot P(H_3) = \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{4}.$$

Summing the three joint probabilities provides the probability of the event E:

$$P(E) = P(H_1 \wedge E) + P(H_2 \wedge E) + P(H_3 \wedge E) = \frac{1+2+3}{12} = \frac{1}{2}.$$

Dividing the joint probabilities by the probability of the event results in the posterior distribution:

$$P(H_1|E) = \frac{P(H_1 \wedge E)}{P(E)} = \frac{1/12}{1/2} = \underline{\underline{\frac{1}{6}}},$$

$$P(H_2|E) = \frac{P(H_2 \wedge E)}{P(E)} = \frac{1/6}{1/2} = \underline{\underline{\frac{1}{3}}},$$

$$P(H_3|E) = \frac{P(H_3 \wedge E)}{P(E)} = \frac{1/4}{1/2} = \underline{\underline{\frac{1}{2}}}.$$

Therefore, after the observation of the coin landing *Heads*, the (posterior) probability of the first hypothesis H_1 has decreased, that of the second hypotheses H_2 stays the same, and the probability of the third hypothesis H_3 has increased.

Let us turn to the final representation of cab problem: the odds format.

4.4.1.6 Bayes-Theorem in Odds Format

The representation of a Bayesian inference problem by means of outcome trees and contingency tables results in a good understanding of the problem structure. Both elucidate the relevance of the different terms in Bayes formula. The odds format of Bayes theorem has two favorable aspects:

1. It simplifies the computation of the posterior probability, and

2. It reveals the important quantities underlying Bayesian reasoning in different contexts, like theories of confirmation (cf. Fitelson, 1999; Hawthorne, 2011) of legal reasoning (Schum, 1994).

The odds format derives directly from Bayes formula by dividing the Bayes formula for computing $P(H|E)$ by the that for computing $P(\bar{H}|E)$, assuming, for the moment, that there are only two relevant hypotheses. In case of the cab problem the two relevant probabilities are: $P(\text{B}|\text{»B«})$ and $P(\text{G}|\text{»B«})$, i.e. the probability of the cab being blue vs green, given the actual testimony of the witness. The relevant expressions are:

$$P(\text{B}|\text{»B«}) = \frac{P(\text{»B«}|\text{B}) \cdot P(\text{B})}{P(\text{»B«})}$$
$$P(\text{G}|\text{»B«}) = \frac{P(\text{»B«}|\text{G}) \cdot P(\text{G})}{P(\text{»B«})}.$$

The symbols have the usual meaning:

B = The cab is blue.
G = The cab is green.
»B« = The witness identifies the cab as »blue«.

By dividing the first expression by the second one the term $P(\text{»B«})$ cancels and one gets:

$$\frac{P(\text{B}|\text{»B«})}{P(\text{G}|\text{»B«})} = \frac{P(\text{»B«}|\text{B})}{P(\text{»B«}|\text{G})} \cdot \frac{P(\text{B})}{P(\text{G})}.$$

This is the odds format representation. It has two benefits:

1. The probability $P(\text{»B«})$ need not be computed.
2. The quotients on the right hand side can be simplified by multiplying numerator and denominator by the same number. Specifically, one can use frequencies instead of

Using the numbers of the cab problem the odds format leads to the following result.

$$\begin{aligned}\frac{P(\text{B}|\text{»B«})}{P(\text{G}|\text{»B«})} &= \frac{P(\text{»B«}|\text{B})}{P(\text{»B«}|\text{G})} \cdot \frac{P(\text{B})}{P(\text{G})} \\ &= \frac{0.8}{0.2} \cdot \frac{0.15}{0.85} \\ &= \underline{\underline{\frac{0.12}{0.17}}}\end{aligned}.$$

Note however, the quotients can also be written by means of frequencies by multiplying the nominator and denumerator by 100:

$$\frac{P\left(\mathrm{B}\,|\,»\mathrm{B}«\right)}{P\left(\mathrm{G}\,|\,»\mathrm{B}«\right)}=\frac{80}{20}\cdot\frac{15}{85}=\frac{12}{\underline{\underline{17}}}.$$

However, the quotients can be further simplified:

$$\frac{P\left(\mathrm{B}\,|\,»\mathrm{B}«\right)}{P\left(\mathrm{G}\,|\,»\mathrm{B}«\right)}=\frac{4}{1}\cdot\frac{3}{17}=\frac{12}{\underline{\underline{17}}}.$$

The possibility to simplify the numerator and denominator is extremely helpful in simplifying the computation.

The resulting term:

$$\frac{P\left(\mathrm{B}\,|\,»\mathrm{B}«\right)}{P\left(\mathrm{G}\,|\,»\mathrm{B}«\right)}=\frac{12}{\underline{\underline{17}}}$$

is called the *posterior odds*. In order to get the required posterior probability one simply divides the posterior odds by 1 plus the posterior odds, in the present case:

$$P\left(\mathrm{B}\,|\,»\mathrm{B}«\right)=\frac{P\left(\mathrm{B}\,|\,»\mathrm{B}«\right)/P\left(\mathrm{G}\,|\,»\mathrm{B}«\right)}{1+P\left(\mathrm{B}\,|\,»\mathrm{B}«\right)/P\left(\mathrm{G}\,|\,»\mathrm{B}«\right)}=\frac{12/17}{1+12/17}=\frac{12}{\underline{\underline{29}}}.$$

However, in case of two hypotheses the result can be found much easier by dividing the numerator of the odds (12) by the sum of the numerator and denominator (12+17 = 29).

The following example illustrates the computational benefits of using odds format.

Ex. 4-37: Bayes theorem in odds format

Given:

Instructor M has two students, S_1 and S_2, who perform their exercises regularly.

Student S_1 is capable of solving 2/3 of the problems whereas Student S_2 solves on average 1/3 of the problems only.

M selects randomly one of the two students using the following procedure: He rolls a fair die and if the number of points is either 1 or 2 he selects S_1, otherwise S_2 is selected.

The student selected receives 5 problems from the pool of exercises. She solves 3 of the 5 problems.

What is the probability that the selected student is S_1?

Comment:

It might be argued that the presented formulation of the problem does not determine a unique probability model. One further assumption is required: The probability of a solution is determined entirely by the ability of the student to solve the problem as given by the probabilities stated above. This precludes the influence of factors other than the student's ability.

In addition, *conditional stochastic independence* of solving the problems is assumed: Given the ability of the student, the probability of solving a given problem is independent of the probability of solving a different problem (this is true for all pairs of problems).
On the basis of the given specification of the problem we can determine the probability that 3 out of 5 problems are solved by student S_i. This probability is given by the binomial distribution:

$$P\left(N=3\middle|S_i\right)=\binom{5}{3}\cdot\pi_i^3\cdot\left(1-\pi_i\right)^2 \quad (i=1,2).$$

The symbols have the following meaning:

π_i denotes the probability that student S_i solves a problem $(i=1,2)$. In the actual case: $\pi_1=2/3$ and $\pi_2=1/3$.

N denotes the number of problems solved.

In order to solve the problem the prior probabilities of the two students due to the selection process are required: $P(S_1)=1/3$ and $P(S_2)=2/3$.

$\binom{5}{3}=\frac{5!}{3!2!}=10$ denotes the binomial coefficient.

Bayes theorem in odds format enables a convenient way to compute the required probability:

$$\frac{P\left(S_1\middle|N=3\right)}{P\left(S_2\middle|N=3\right)}=\frac{(2/3)^3\cdot(1/3)^2}{(1/3)^3\cdot(2/3)^2}\cdot\frac{(1/3)}{(2/3)}=1$$

Consequently, $P\left(S_1\middle|N=3\right)=P\left(S_2\middle|N=3\right)=1/2$.

Note that the computation using odds format is simplified since the binomial coefficient:

$$\binom{5}{3}=\frac{5!}{3!2!},$$

is part of the numerator and denominator and thus cancels.
Assume, for the moment, that the process of selection had been performed by throwing a fair coin. In the this case the prior probabilities would have been the same:

$$P(S_1)=P(S_2)=1/2.$$

The would have led to the following result:

$$\frac{P\left(S_1\middle|N=3\right)}{P\left(S_2\middle|N=3\right)}=\frac{(2/3)^3\cdot(1/3)^2}{(1/3)^3\cdot(2/3)^2}\cdot\frac{(1/2)}{(1/2)}=\frac{2}{1}.$$

In this case, the resulting probability that S_1 had been selected was: $P(S_1|N=3)=2/3$.

Ex. 4-37 demonstrates strikingly the advantage of the odds format for computing posterior probabilities. This simplification is due to the two characteristics of the odds format presented above:

1. The probability of the conditioning event need not be computed, in the present case this probability is given by:

$$P(N=3)=P(N=3|S_1)\cdot P(S_1)+P(N=3|S_2)\cdot P(S_2)$$
$$=\binom{5}{3}\cdot(2/3)^3\cdot(1/3)^2\cdot(1/3)+\binom{5}{3}\cdot(1/3)^3\cdot(2/3)^2\cdot(2/3)$$

2. The numerators and denominators of the quotients may be simplified. In the present case the binomial coefficient can be ignored.

The odds format of Bayes theorem can also be used in case of more than two hypotheses.

Concept 4-7: *Bayes theorem in odds format*

Given:

- A set $H_1, H_2, \ldots, H_n$ of exclusive and exhaustive hypotheses with a priori probabilities: $P(H_1), P(H_2), \ldots, P(H_n)$.
- The likelihoods $P(E|H_1), P(E|H_2), \ldots, P(E|H_n)$, i.e. the probabilities of the observed evidence E given the hypotheses.

Bayes theorem in odds format is used to calculate the posterior odds:

$$\omega_{i,n}=\frac{P(H_i|E)}{P(H_n|E)}, \qquad (i=1,2,\ldots,n-1)$$

where hypothesis H_n is used as a reference. The final result is independent of which hypothesis is taken as the reference. However, each of the posterior odds has to be computed using the same reference hypothesis:

$$\frac{P(H_i|E)}{P(H_n|E)}=\frac{P(E|H_i)}{P(E|H_n)}\cdot\frac{P(H_i)}{P(H_n)} \qquad (i=1,2,\ldots,n-1)$$

In abbreviated form:

$$\omega_{i,n}=\lambda_{i,n}\cdot\pi_{i,n}.$$

The symbols have the following names:

$$\omega_{i,n} = \frac{P(H_i|E)}{P(H_n|E)}$$ are the *posterior odds*,

$$\lambda_{i,n} = \frac{P(E|H_i)}{P(E|H_n)}$$ are called the *likelihood ratios*, and

$$\pi_{i,n} = \frac{P(H_i)}{P(H_n)}$$ are the *prior odds*.

Comment:

The computation of the odds with more than two hypotheses benefits from the same possible simplifications as in case of two hypotheses (see above).

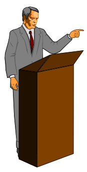
On the basis of the computed posterior odds the posterior distribution is calculated by simply dividing each of the posterior odds by the sum of the posterior odds.

Note that one of the *n* posterior odds has the value 1.0, in the actual case:

$$\omega_{n,n} = \frac{P(H_n|E)}{P(H_n|E)} = 1.$$

Thus, the posterior probability is given by the equation:

$$P(H_i|E) = \frac{\omega_{i,n}}{1 + \sum_{j=1}^{n-1} \omega_{j,n}}.$$



Comment 4-6: On the usage of odds format

According to my experience students are at first deterred by the odds format, and refuse to use it since they do not realize the computational simplification that the format provides. This was also the case when I first encountered Bayes theorem in odds format.

To illustrate the computations using odds format with more than two hypotheses let us redo Ex. 4-36 (on page 171).

Ex. 4-38: Bayesian reasoning in odds format:

Given: Three hypotheses concerning the bias of a coin:

$$H_1 : \pi_1 = 1/4$$

$$H_2 : \pi_2 = 1/2$$

$$H_3 : \pi_3 = 3/4$$

Where π_i denotes the probability of the coin landing *Heads*, under the different hypotheses H_i $(i = 1, 2, 3)$:

$$P\left(Heads \middle| H_i\right) = \pi_i \text{ and } P\left(Tails \middle| H_i\right) = 1 - \pi_i.$$

It is assumed that the three hypotheses represent the only possibilities, i.e. they exhaust the space of hypotheses.
In addition, assume that the prior probabilities for each of the three hypotheses are the same, i.e.:

$$P(H_1) = P(H_2) = P(H_3) = 1/3.$$

The coin is tossed and lands *Heads* (*H*). What is the probability of the three hypotheses given the observed outcome *E*: $P\left(H_i \middle| E\right)$, ($i$ = 1, 2, 3).

We use H_3 as the reverence hypothesis. The two posterior odds are thus given by:

$$\frac{P\left(H_1 \middle| E\right)}{P\left(H_3 \middle| E\right)} = \frac{1/4}{3/4} \cdot \frac{1/3}{1/3} = \frac{1}{3} \text{ and } \frac{P\left(H_2 \middle| E\right)}{P\left(H_3 \middle| E\right)} = \frac{2/4}{3/4} \cdot \frac{1/3}{1/3} = \frac{2}{3}$$

Comments:
Since the prior probabilities of the three hypotheses are identical the prior odds are always 1 and can thus be ignored.

Using the computed odds the posterior distribution is given by:

$$P\left(H_1 \middle| E\right) = \frac{P\left(H_1 \middle| E\right) / P\left(H_3 \middle| E\right)}{1 + P\left(H_1 \middle| E\right) / P\left(H_3 \middle| E\right) + P\left(H_2 \middle| E\right) / P\left(H_3 \middle| E\right)} = \frac{1/3}{1 + 1/3 + 2/3} = \underline{\underline{\frac{1}{6}}}$$

$$P\left(H_2 \middle| E\right) = \frac{P\left(H_2 \middle| E\right) / P\left(H_3 \middle| E\right)}{1 + P\left(H_1 \middle| E\right) / P\left(H_3 \middle| E\right) + P\left(H_2 \middle| E\right) / P\left(H_3 \middle| E\right)} = \frac{2/3}{1 + 1/3 + 2/3} = \underline{\underline{\frac{1}{3}}}$$

$$P\left(H_3 \middle| E\right) = \frac{P\left(H_3 \middle| E\right) / P\left(H_3 \middle| E\right)}{1 + P\left(H_1 \middle| E\right) / P\left(H_3 \middle| E\right) + P\left(H_2 \middle| E\right) / P\left(H_3 \middle| E\right)} = \frac{1}{1 + 1/3 + 2/3} = \underline{\underline{\frac{1}{2}}}$$

Comment: It would be helpful to perform and compare the computations with and without odds format (cf. Ex. 4-36, on page 171).

As noted previously, Bayes theorem in odds format provides new insights with respect to the interpretation of the involved quantities, i.e the odds. To better understand the significance of the notion of *odds* the concept of (fair) betting odds has to be elucidated first.

Concept 4-8: *Fair betting odds and probability of winning*

Odds are used in betting situation to describe the relative changes of winning vs. losing. For example,

The odds that YB Bern will win the Swiss soccer championship next year are 2:1.

This statement asserts that the probability of YB Bern winning the championship is $2/3$ whereas the probability of YB Bern not winning is $1/3$.

The *betting odds* are thus simple the quotient of these two probabilities:

$$\omega_b = \frac{P(winning)}{1-P(winning)},$$

where ω_b denotes the betting odds.

The betting odds determine the devision of the stack of a game, i.e. how the money has to be divided. If the stack is devided according to the odds one speaks of *fair betting odds* since, in this case, the expected loss of the two bettors betting on the two opposite options (winning vs. losing) is zero. This can be demonstrated in a straightforward way.

Given:

S = the whole stake, i.e. the total amount of money put into the game by both players. For the sake of concreteness, assume that the the whole stake is S = Sfr. 20.-

p = the probability of a favorable outcome. In our example, assume that the probability is $p = 2/3$ that YB Bern will win the championship.

The game is fair, if the player betting on YB Bern's winning has payed Sfr. $p \cdot S$ and her adversary has put in Sfr. $(1-p) \cdot S$.

A simple calculation elucidates that the the expected outcome $E(\text{outcome})$ is zero (for both players). For, example, the expected outcome for the player betting in favor of YB Bern is:

$$E(\text{outcome}) = p \cdot (S - p \cdot S) + (1-p) \cdot (-p \cdot S) = 0$$

The term $S - p \cdot S$ represents the net win in case of a favorable outcome (that occurs with probability p), whereas the term $-p \cdot S$ represents the loss in case of YB Bern not winning the champion ship (namely the stack of the player).

Since neither the bettor nor her adversary can expect a win or loss the bet is fair.

Let us now return to Bayes theorem in odds format. The relevant odds involved are:

1. The prior odds $P(H_i)/P(H_n)$ represent the relative strength of the belief in hypothesis H_i compared to H_n prior to the relevant observation(s).

   For the cab problem the prior odds are $3/17$ in favor of a blue taxi. This means that, without further information, the chances of a blue taxi instead of a green one having caused the accident are 3 to 17. Consequently in a fair bet with a total stake of Sfr. 20.- a bettor would be willing to risk Sfr. 3.- in order to get Sfr. 17.- (the bettor wins Sfr. 17.- in addition to her personal stake of Sfr. 3.- if a blue cab was actually involved in the accident, and she loses her stake in case of a green cab being concerned).

Note that the exact numbers of blue and green cabs in town are not relevant, only the odds count. This fact is also reflected by the sampling version of the cab problem where only the prior odds were relevant for sampling the cabs and not the exact numbers (cf. Section 4.4.1.3).

2. The second quantity involved in the odds format of Bayes theorem is the likelihood ratio $LR = P(E|H_i)/P(E|H_n)$. It represents the degree to which the observed data favor one of the two hypotheses. The following relationship holds.

$$LR = \begin{cases} >1 & \Leftrightarrow \quad \text{the data favor } H_i \text{ over } H_n \\ \ \ 1 & \Leftrightarrow \quad \text{the data equally favor both hypotheses} \\ <1 & \Leftrightarrow \quad \text{the data favor } H_n \text{ over } H_i \end{cases}$$

Thus, the likelihood ratio represents the impact of the (new) evidence E on the evaluation of the relative strength of the hypothesis. It determines how strongly the prior odds are revised due to evidence given:

(a) If $LR > 1$ the posterior odds will be greater than the prior odds. Consequently, the relative strength of the belief in hypothesis H_i compared to H_n will be increased.

(b) If $LR = 1$ the posterior odds will be identical to the prior odds. Consequently, the relative strength of the belief in hypothesis H_i compared to H_n will remain the same.

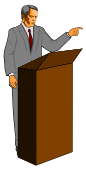
(c) If $LR < 1$ the posterior odds will be lower than the prior odds. Consequently, the relative strength of the belief in hypothesis H_i compared to H_n will be decreased.

In case of the cab problem $LR = 4$ (in favor of the blue cab) thus the prior odds of $3/17$ are increased by a factor of 4 resulting in the posterior odds of $12/17$. Thus, in case of a fair bet, our bettor in favor of a blue cab has now to risk Sfr. 12.- in order to get Sfr. 17.-.

With odds format it is quite easy to judge whether, in case of two hypotheses being involved, the posterior probability of hypothesis H is greater than 0.5. This is the case if the posterior odds of H vs. $\overline{H}$ are greater than 1.0.

In addition, odds format enables a quick estimation of whether the prior »bias« in favor of one hypothesis can be traded off by the likelihood ratio. For example, for the cab problem the prior bias in

favor of the green cabs is $17/3$ and the likelihood ratio is 4 (in favor of the blue cabs). By consequence, the prior bias cannot be offset by the likelihood ratio since $17/3 > 4$. This leads immediately to the conclusion that the posterior probability of a blue taxi having caused the accident will be smaller than 50% since the posterior odds are lower than 1.0.
Finally, betting odds have an additional advantage over probabilities: They enable a more convenient way to quantify uncertainty since they do not require normalizing conditions of probabilities (i.e. the probabilities of exhaustive and exclusive events have to sum to 1.0). It turned out that the usage of odds resulted in more consistent inferences (Slovic & Lichtenstein, 1971).



Comment 4-7: Comprehension of odds in different cultures:

According to my personal experience there exists a sharp distinction between Germans and people from Anglo-Saxon countries with respect to the employment of betting odds: The latter are much more common in the Anglo-Saxon countries (specifically in Great Britain) whereas Germans usually exhibit a low level of comprehension of odds. This is probably due to different betting habits.

We have discussed different representations of simple Bayesian problems as well as different methods on how to solve them. As already noted above, Bayesian reasoning is but a special case of probabilistic reasoning. In the following section, the general structur of probabilistic reasoning will be exhibited.

4.4.2 The Structure of Probabilistic Reasoning

In order to understand probabilistic reasoning one has to consider three different *types of probabilities*:

1. Joint probabilities,
2. Marginal probabilities, and
3. Conditional probabilities.

Each of these probabilities has already been encountered in the previous chapters. However, for the sake of convenience, these quantities will be reviewed here once again. In addition, there exist three *probabilistic operations*:

1. Combination of probabilistic information,
2. Marginalization, and
3. Conditioning.

Probabilistic reasoning consists in applying these operations in order to »move« between the three types of probabilities presented above. The operations map the different types of probabilities onto another one. Bayesian reasoning involves each of the three operations. Let us now discuss the different types of probabilities as well as the probabilistic operations in detail.

4.4.2.1 The Different Types of Probabilities

In the following we assume that there exists a set of random variables: $X_1, X_2, \ldots, X_n$ (Remember that a random variable is variable that takes on different values with given probabilities).

***Concept 4-9:** Joint distribution*

The joint probability distribution $P(X_1, X_2, \ldots, X_n)$ represents the probability of each possible combination of the values of the random variables.

Notation 4-5: Joint probabilities

1. The symbol $P(X_1, X_2, \ldots, X_n)$ represents a whole table of joint probabilities (cf. Notation 4-2 and Notation 4-3 on page 132 and 133, respectively).
   Each entry of table represents the probability of a specific event $P(X_1 = x_{1i}, X_2 = x_{2j},, \ldots, X_n = x_{nk})$ that is given by the fact that variable X_1 takes on the value x_{1i}, X_2 takes on the value x_{2j}, and so on.
2. Instead of using $P(X_1 = x_{1i}, X_2 = x_{2j},, \ldots, X_n = x_{nk})$ the simpler notation $P(x_{1i}, x_{2j},, \ldots, x_{nk})$ is used for denoting the joint probability of the variables taking a specific combination of values.

Examples of joint probabilities have been provided above. For example, in Section 4.3.1.3 the following probability has been considered:

$$P(\text{lung cancer, high blood pressure, male, smoker, age} \geq 50).$$

It denotes the probability that a person in the population is a male smoker over 50 and has a lung cancer as well as a high blood pressure.

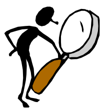
The following principle concerning the joint distribution is of focal importance.

***Principle 4-4:** Joint distribution and probability information:*

The joint distribution contains the complete probability information about the underlying random variables.
By consequence, any question concerning probabilistic information about the underlying events can be answered by reference to the joint distribution.

Due to the favorable characteristic of completeness of the joint distribution it seems to be the optimal representation of probability information. However, this is not necessarily the case for the following reasons:

1. In case of many random variables the joint distribution becomes too complex. Assume, for example, 100 random variables with two outcomes each. In this case the joint distribution comprises $2^{100} = 1.27 \times 10^{30}$ probabilities. Assume that only one bit is required for storing a single probability. An exabyte comprises 8×10^{18} bits. Thus, to store the whole table 100 billion exabytes (about 10^{11}) are needed. Estimates assume that, at present (2016), the whole disk space of all computers on earth comprises about 2500 exabytes.
2. The joint distribution does not reveal important information. For example, it does not reveal whether two or more events are stochastically dependent or not (cf. Section 4.3.1.4).



Concept 4-10*: Marginal distribution*

Given:

A joint distribution $P(X_1, X_2, \ldots, X_n)$ on the set of random variables $X_1, X_2, \ldots, X_n$.

The *marginal distribution* with respect to the given set of variables is a probability distribution on a proper subset of the given random variables.

For example, the distributions $P(X_1)$, $P(X_2)$, …, $P(X_n)$ or $P(X_1, X_2)$ are marginal distributions with respect ot the given set of variables.

Given our example, above:

$$P(\text{lung cancer, high blood pressure, male, smoker, age} \geq 50).$$

The probabilities:

$$P(\text{lung cancer})$$

$$P(\text{high blood pressure, male, smoker}), \text{ and}$$

$$P(\text{lung cancer, high blood pressure, male, smoker})$$

are examples of marginal probabilities with respect to the given joint probability.

The third type of probabilities, conditional probabilities, has been explained in great detail above in Chapter 4.3.1, and the explication will not be repeated here. Let us now turn to the probabilistic operations.

4.4.2.2 Probabilistic Operations

The first operation, the *combination of probability information*, consists in the generation of a joint probability distribution by combining conditional and marginal probabilities.

The combination is based on the equation of the conditional distribution discussed in detail in Chapter 4.3.1.

$$P\left(X_1, X_2, \ldots, X_k \middle| X_{k+1}, X_{k+2}, \ldots, X_n\right) = \frac{P\left(X_1, X_2, \ldots, X_n\right)}{P\left(X_{k+1}, X_{k+2}, \ldots, X_n\right)}$$

Multiplying of both sides with the denominator on the right hand side of the equation results in (exchanging sides):

$$P\left(X_1, X_2, \ldots, X_n\right) = P\left(X_1, X_2, \ldots, X_k \middle| X_{k+1}, X_{k+2}, \ldots, X_n\right) \cdot P\left(X_{k+1}, X_{k+2}, \ldots, X_n\right)$$

For example, with only two variable *X* and *Y*, we have the familiar equation: $P(X,Y) = P(X|Y) \cdot P(Y)$.

Repeated application of the basic equation enables one to derive the frequently used *chain rule*. We restrict ourselves to 4 variables. The generalization to arbitrary many variables is straightforward.

Ex.4-39: Chain rule

The chain rule with four variables *X*, *Y*, *Z*, and *W* looks like this:

$$P(X,Y,Z,W) = P(X|Y,Z,W) \cdot P(Y|Z,W) \cdot P(Z|W) \cdot P(W)$$

The chain rule results from repeated application of the basic equation for computing the joint probability:

$$P(X,Y,Z,W) = P(X|Y,Z,W) \cdot P(Y,Z,W)$$

$$P(Y,Z,W) = P(Y|Z,W) \cdot P(Z,W)$$

$$P(Z,W) = P(Z|W) \cdot P(W)$$

Successive substitution of the right hand sides results in the chain rule as given above.

We come to the discussion of the second operation. The operation of *marginalization* consists in a summation performed on the joint distribution. The result of marginalization is a marginal distribution. The summation is taken over all combination of values of those variables that do not make up the resulting marginal distribution (These variables are »summed out« of the distribution). Formally, the operation of marginalization can be represented as follows:

$$P\left(X_1, X_2, \ldots, X_k\right) = \sum_{\substack{\text{all combinations of} \\ \text{values of the variables} \\ X_{k+1}, X_{k+2}, \ldots, X_n}} P\left(X_1, X_2, \ldots, X_k, X_{k+1}, X_{k+2}, \ldots, X_n\right)$$

The sum in the equation runs over all combinations of values $\left(x_{k+1,i}, x_{k+2,j}, \ldots, x_{nv}\right)$ of the variables $X_{k+1}, X_{k+2}, \ldots, X_n$. In the resulting marginal distribution these variables are no longer present.

Notation 4-6: Marginalization

The operation of marginalization is represented more succintly as follows:

$$P(X_1, X_2, \ldots, X_k) = \sum_{X_{k+1}, X_{k+2}, \ldots, X_n} P(X_1, X_2, \ldots, X_k, X_{k+1}, X_{k+2}, \ldots, X_n)$$

This notation stresses the fact that the sum runs over the variables:

$$X_{k+1}, X_{k+2}, \ldots, X_n,$$

that are »summed out« and are no longer part of the resulting marginial distribution.

In case of continuous variables summation is replaced by integration. The final probabilitistic operation, conditioning, has been discussed in in detail in Chapter 4.3.1. Let us now take a look at Bayesian reasoning from the perspective of probabilitistic operations.

4.4.2.3 Bayesian Computations and Probabilistic Operations

Figure 4-13 illustrates the conception of Bayesian inference in terms of probabilistic operations for the cab problem. The distribution of the colors of the cabs $P(Color)$, as well as the conditional distribution of the testimony of the witness, given the color of the cabs $P(Witness|Color)$ are provided in the description of the problem. Bayesian reasoning involves the application of the three steps:

- ❑ In the *first step* the given different pieces of information consisting in the conditional and marginal probabilities are combined resulting in the joint distribution. The combination is performed by means of multiplying the different probabilities (application of the chain rule).
- ❑ In the *second step* the marginal distribution is computed from the joint distribution by means of marginalization: summation over variables that do not make up the marginal distribution.
- ❑ In the *third step* the joint and marginal distribution is used to compute the conditional distribution using the definition of conditional probabilities. The concrete operation consists in dividing the joint probabilities by the marginal probabilities.

In Section 4.4.1 we discussed different methods for performing the probabilisitic operations. Specifically, outcome trees can be used to implement the combination of information to get the relevant joint probablilites $P(\text{Blue, »Blue«})$ and $P(\text{Green, »Blue«})$. The joint probabilities or joint frequencies are found at the leaves of the outcome tree (Section 4.4.1.2). Similarly, a contingency table may be constructed with the entries of the table containing the joint probabilities and frequencies, resprectively (Section 4.4.1.4). Finally, the Bayes formula is implements the chain rule to compute the required joint probabilities (Section 4.4.1.5).

The process of marginalization that consists in summing of the relevant joint probabilities (summing over the colors of the cabs):

$$P(\text{»Blue«}) = P(\text{Blue, »Blue«}) + P(\text{Green, »Blue«}),$$

is most conveniently performed in the context of contingency tables by summing the entries to get the margins of the table. However, it was also demonstrated how this operation can also be performed in the context of outcome trees (cf. Figure 4-10and Figure 4-8) or by using Venn diagrams (Figure 4-9).

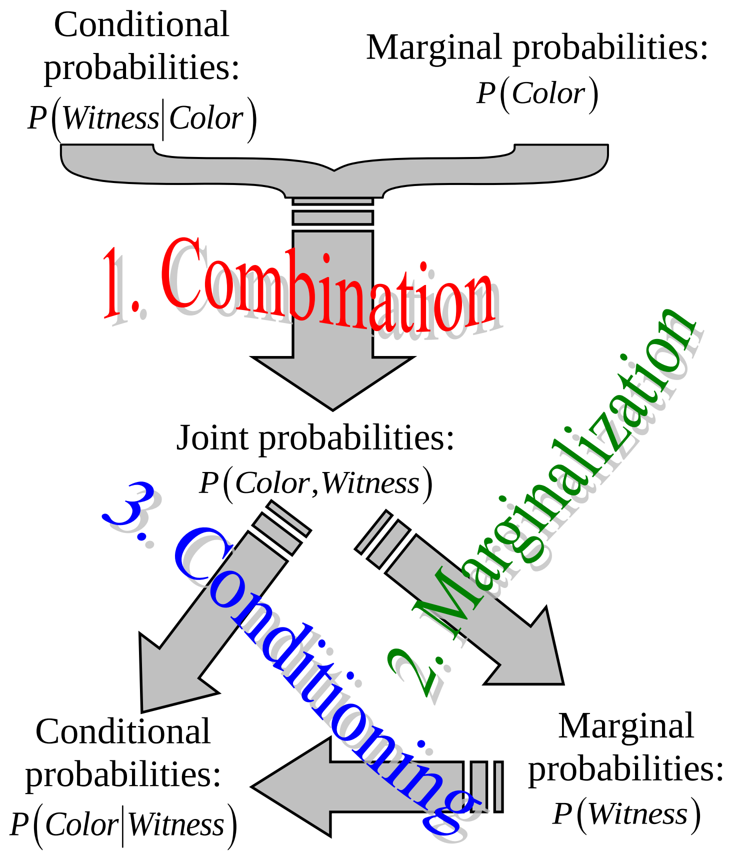


Figure 4-13: *Basic operations in Bayesian reasoning for the cab problem*

Finally, the process of conditioning that requires the division of the relevant joint probability by the relevant marginal probability,

$$P(\text{Blue}|\text{»Blue«}) = \frac{P(\text{Blue, »Blue«})}{P(\text{»Blue«})},$$

has also demonstrated above, for the different methods of representing the cab problem.

To illustrate the different probabilities and operations we use an extension of the cab problem, and demonstrate how the combination, marginalization and conditioning can be performed using outcome trees, contingency tables and Bayes formula.

Ex. 4-40: The Extended Cab Problem:
A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85% of the cabs in the city are Green (*G*) and 15% are Blue (*B*).

(b) A witness identified the cab as »blue«. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

(c) A second witness identified the cab as »green«. The court also tested the reliability of this witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified the color of the blue cabs 60% of the time and failed 40% of the time. He correctly identified the color of the green cabs 70% of the time and failed 30% of the time.

What is the probability that the cab involved in the accident was blue rather than green?

Notation:
The following symbols are used to denote different events:

B = The cab is blue.
G = The cab is green.
»B_1« = The first witness identified the cab as »blue«.
»G_1« = The first witness identified the cab as »green«.
»B_2« = The second witness identified the cab as »blue«.
»G_2« = The second witness identified the cab as »green«.

The required probability is: $P\left(B \middle| »B_1«, »G_2«\right)$.

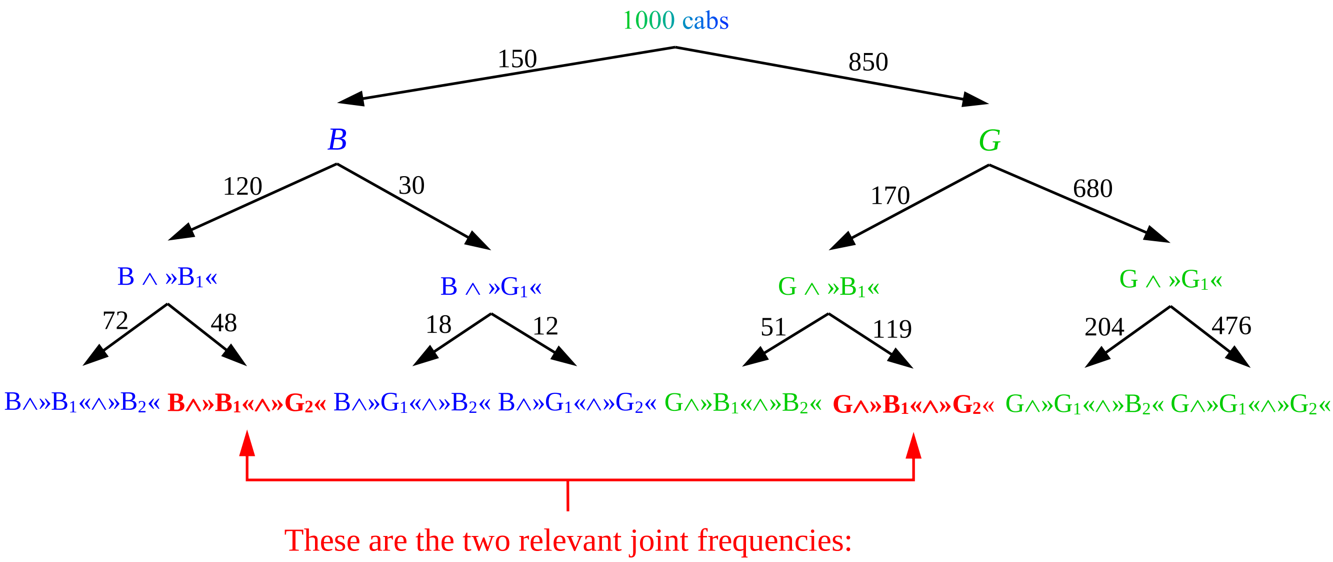
Comment:
Contrary to the orginal version with two variables, the extended version of the cab problem involves three variables:

(a) The color of the cabs;
(b) The testimony of the first witness;
(c) The testimony of the second witness.

Let us now demonstrate the probablisitic operations using outcome trees in frequency format.

4.4.2.3.1 Probabilitistic Resoning with Outcome Trees

Figure 4-14 depicts the outcome tree representation of the extended cab problem. The tree contains an additional layer that represents the joint events resulting from the combinations of the possible values of the three variables (at the leaves of the tree). The numbers on the arrows pointing to the leaves represent the joint frequencies.



$$P\left(B \middle| »B_1«,»G_2«\right) = \frac{\#\left(B,»B_1«,»G_2«\right)}{\#\left(B,»B_1«,»G_2«\right) + \#\left(G,»B_1«,»G_2«\right)} = \frac{\#\left(B,»B_1«,»G_2«\right)}{\#\left(»B_1«,»G_2«\right)}$$

$$= \frac{48}{48+119} = \underline{\underline{\frac{48}{167}}}$$

Figure 4-14: *Representation of the extended cab problem by means of an outcome tree in frequency format.*

These numbers result form multiplying the joint frequencies of the two variables representing the color of the cab and the identification of the first witness by the conditional probability of the testimony of the second witness given the combination of values of the other two variables. For example:

$$\#\left(B,»B_1«,»G_2«\right) = \#\left(B,»B_1«\right) \cdot P\left(»G_2« \middle| B,»B_1«\right) = 120 \cdot 0.4 = \underline{\underline{48}}.$$

In the lower part of Figure 4-14 the operations of marginalization and conditioning are shown. The process of marginalizaton consists in summing the two relevant joint frequencies:

$$\underbrace{\#\left(B,»B_1«,»G_2«\right)}_{48} + \underbrace{\#\left(G,»B_1«,»G_2«\right)}_{119} = \underbrace{\#\left(»B_1«,»G_2«\right)}_{167}.$$

The operation of conditioning consists in dividing the relevant joint probabibily by this sum:

$$P\left(B \middle| »B_1«,»G_2«\right) = \frac{\#\left(B,»B_1«,»G_2«\right)}{\#\left(»B_1«,»G_2«\right)} = \frac{48}{167}.$$

The outcome tree based on probabilities of frequencies is constructed in the same way with the probabilistic operations being applied to probabilities instead of frequencies (cf. Exercise 4-15).
Let us now represent the problem by means of a contingency table.

4.4.2.3.2 Probabilitistic Resoning with Contingency Tables

We build up the contingency table in two steps that mirrors the construction of the outcome tree. In the first step the contingency table of the joint frequencies involving the color of the cab and the identification of the first witness is constructed. We start with the marginal frequencies of the colors of the cab (150, 850) and split each according to the given conditional probabilities of the answers of the first witness. For example, the 150 blue cabs are split into 120 and 30 since the witness is 80% correct and fails in 20% of the cases. Tab. 4-10 shows the resulting table.

Tab. 4-10: *Contingency table representing joint and marginal frequencies of the colors of the cab and the testimony of the first witness.*

	Classification of the first witness		
Color of the cab	»B_1«	»G_1«	Σ
B	120	30	150
G	170	680	850
Σ	290	710	1000

In the second step the cells of the table are further split according to the conditional probabilities of the identification of the second witness given the colors of the cabs and the answers of the first witness. For example the 120 cases of »blue« identifications of the first witness of blue cabs are split into the number 72 and 48 since 60% of the blue cabs are identified as »blue« by the second witness and the remaining 40% are identified as »green« (60% of 120 is 72). Tab. 4-11 exhibits the resulting table.

Thus the entries of Tab. 4-11 represent the joint frequency distribution of the three variables (color of the cab, answer of the first and second witness).

Tab. 4-11: *Contingency table representing joint and marginal frequencies of the colors of the cab and the testimonies of both witnesses.*

	Classification of the witnesses				
Color	»B_1«,»B_2«	»B_1«,»G_2«	»G_1«,»B_2«	»G_1«,»G_2«	Σ
B	72	48	18	12	150
G	51	119	204	476	850
Σ	123	167	222	488	1000

The last row in Tab. 4-11 contains the marginal distribution of answers of the two witnesses (summed over the color of the cab), i.e., the result of the operation of marginalization.

Division of the entries of the table by these marginal frequencies results in the table of conditional probabilities of the colors of the cabs for the different answers of the witnesses. Tab. 4-12 shows the result.

Tab. 4-12: *Table representing the conditional probabilities of the colors of the cabs given the identifications of the witnesses.*

	Classification of the witnesses			
Color	»B_1«,»B_2«	»B_1«,»G_2«	»G_1«,»B_2«	»G_1«,»G_2«
B	72/123	**48/167**	18/222	12/488
G	51/123	119/167	204/222	476/488
Σ	1	1	1	1

Since it is known that the answers of the witnesses were »B_1« and »G_2« and we are interested in the probability that the cab is blue given this condition, only the respective entry (in red) is relevant for solving the problem.

The determination of the joint frequencies and probabilities, respectively, appears to be straight foreward and easier with outcome trees compared to contingency tables. However, as soon as the joint frequencies (probabilities) are given, the contingency table format is convenient for computing different probabilities, as the following example demonstates.

Ex. 4-41: Computation of various probabilities using contingency tables:

Given:

The joint probabilities of the extended cab problem (cf. Tab. 4-13).

Required: The following probabilities:

- ❑ $P\left(»B_1«,»G_2«\right)$
- ❑ $P\left(»B_1«\right)$

- ❑ $P(»G_2«)$
- ❑ $P(»B_1«|»G_2«)$
- ❑ $P(»G_2«|»B_1«)$

These probabilities can be easily determined using the joint probabilities in the contingency table:

$$P(»B_1«,»G_2«)=\frac{167}{1000} \qquad \text{[last row of the table]}$$

$$P(»B_1«)=P(»B_1«,»B_2«)+P(»B_1«,»G_2«)$$
$$=\frac{123}{1000}+\frac{167}{1000}=\frac{290}{1000} \qquad \text{[marginalization]}$$

$$P(»G_2«)=P(»B_1«,»G_2«)+P(»G_1«,»G_2«)$$
$$=\frac{167}{1000}+\frac{488}{1000}=\frac{655}{1000} \qquad \text{[marginalization]}$$

$$P(»B_1«|»G_2«)=\frac{P(»B_1«,»G_2«)}{P(»G_2«)} \qquad \text{[conditioning]}$$
$$=\frac{167/1000}{655/1000}=\frac{167}{655}$$

$$P(»G_2«|»B_1«)=\frac{P(»B_1«,»G_2«)}{P(»B_1«)} \qquad \text{[conditioning]}$$
$$=\frac{167/1000}{190/1000}=\frac{167}{290}$$

Tab. 4-13: *Contingency table representing joint and marginal probabilities of the colors of the cab and the testimonies of the two witnesses for the extended cab problem.*

	Classification of the witnesses				
Color	»B1«, »B2«	»B1«, »G2«	»G1«, »B2«	»G1«, »G2«	Σ
B	72/1000	48/1000	18/1000	12/1000	150/1000
G	51/1000	119/1000	204/1000	476/1000	850/1000
Σ	123/1000	167/1000	222/1000	488/1000	1

This suggests a combination of the two methods for computing required probabilities:

(a) Determine the joint frequencies or probabilities by means of outcome trees.

(b) Assemble the joint probabilities in a contingency table such that they can be used conveniently for performing the operations of marginalization and conditioning.

Let us next investigate how Bayes formula integrates the probabilisitic operations for the extended cab problem.

4.4.2.3.3 Probabilitistic Operations Embedded in Bayes Formula

As usual the starting point is the definition of the conditional probability that incorporates the process of conditioning.

$$P\left(B\middle|»B_1«,»G_2«\right)=\frac{P\left(B,»B_1«,»G_2«\right)}{P\left(»B_1«,»G_2«\right)}$$

Since the terms on the right-hand side are not given directly they have to be computed using the chain rule for combining the information.

$$\begin{aligned}P\left(B,»B_1«,»G_2«\right)&=P\left(B\right)\cdot P\left(»B_1«\middle|B\right)\cdot P\left(»G_2«\middle|B,»B_1«\right)\\&=0.15\cdot 0.80\cdot 0.60\end{aligned}.$$

Similarly,

$$\begin{aligned}P\left(G,»B_1«,»G_2«\right)&=P\left(G\right)\cdot P\left(»B_1«\middle|G\right)\cdot P\left(»G_2«\middle|G,»B_1«\right)\\&=0.85\cdot 0.20\cdot 0.70\end{aligned}.$$

The denomination of the above equation can be obtained by applying the operation of marginalization which consists in adding these two joint probabilities:

$$P\left(»B_1«,»G_2«\right)=P\left(B,»B_1«,»G_2«\right)+P\left(G,»B_1«,»G_2«\right)$$

Putting these steps together results in Bayes formula:

$$P\left(B\middle|»B_1«,»G_2«\right)=\frac{P\left(B\right)\cdot P\left(»B_1«\middle|B\right)\cdot P\left(»G_2«\middle|B,»B_1«\right)}{P\left(B\right)\cdot P\left(»B_1«\middle|B\right)\cdot P\left(»G_2«\middle|B,»B_1«\right)+P\left(G\right)\cdot P\left(»B_1«\middle|G\right)\cdot P\left(»G_2«\middle|G,»B_1«\right)}$$

This illustrates how Bayes formula integrates the three probabilistic operations.

4.4.2.3.4 The Representation of Conditional Independence with Different Formats

The representation and solution of the extended cab problem incorporates a hidden assumption: Given the color of the cab, the testimonies of the two witnesses are stochastically independent, in symbols:

$$P\left(W_1\middle|C,W_2\right)=P\left(W_1\middle|C\right).$$

W_i denotes the testimony of Witness i (i = 1, 2), and C symbolizes the color of the cab. Thus, there exists conditional stochastic independence between the testimonies of the two witnesses, given the color of the cab (For a general discussion of the concepts of stochastic independence and conditional stochastic independence, cf. Chapter 4.3.1.4).

In the following two issues will be discussed:

(a) Why is the assumption of conditional stochastic independence plausible for the given example, whereas the assumption of unconditional stochastic independence between the two testimonies is implausible?

(b) How is conditional independence revealed by the different representations of the problem?

Let us first demonstrate that the conditional independence in question holds whereas unconditional independence does not exist. Unconditional independence between the testimonies holds, if the marginal probability of the testimony of one witness is identical to the conditional probability of the testimony given the testimony of the other witness: $P(W_1|W_2)=P(W_1)$ or, equivalently, $P(W_2|W_1)=P(W_2)$. In Ex. 4-41 it was shown that:

$P(»B_1«)=290/1000$, and

$P(»B_1«|»G_2«)=167/655$.

Since the two probabilities are not the same, unconditional independence does not hold.
Now, let us show that unconditional independence hold, i.e. the conditional probability of the testimony of a witness given the color of the cab is the same as the conditional probability of the testimony given the color of the cab and the testimony of the other witness. This has to be shown for all combination of values. Using the data from Tab. 4-11, on page 191 as well as the results in Ex. 4-41 we get:

$$P(»B_1«|B)=\frac{P(B,»B_1«)}{P(B)}=\frac{P(B,»B_1«,»B_2«)+P(B,»B_1«,»B_2«)}{P(B)},$$
$$=\frac{48+72}{150}=\underline{\underline{\frac{4}{5}}}$$

$$P(»B_1«|B,»B_2«)=\frac{P(B,»B_1«,»B_2«)}{P(B,»B_2«)}=\frac{P(B,»B_1«,»B_2«)}{P(B,»B_1«,»B_2«)+P(B,»G_1«,»B_2«)},$$
$$=\frac{72}{72+18}=\underline{\underline{\frac{4}{5}}}$$

$$P(»B_1«|B,»G_2«)=\frac{P(B,»B_1«,»G_2«)}{P(B,»G_2«)}=\frac{P(B,»B_1«,»G_2«)}{P(B,»B_1«,»G_2«)+P(B,»G_1«,»G_2«)}.$$
$$=\frac{48}{48+12}=\underline{\underline{\frac{4}{5}}}$$

In the same way it can be shown that (Exercise 4-17):

$P(»B_1«|G)=P(»B_1«|G,»B_2«)=P(»B_1«|G,»G_2«)$.

Note also that:

$P(»G_1«|B)=1-P(»B_1«|B)$,

$P(»G_1«|B,»B_2«)=1-P(»B_1«|B,»B_2«)$, and

$P(»G_1«|B,»G_2«)=1-P(»B_1«|B,»G_2«)$.

Thus the given identities indicate that the conditional independence in question holds. Let us return to the two issues stated above:

(a) The conditional independence of the testimony given the colors means that the testimony of one witness has no direct influence on that of the other one. This is plausible if the witnesses do not know of each other, and, by consequence, the probability of naming a specific color is influenced only by the color of the cab and the reliability in identifying the correct color under the given condition but not by the testimony of the other witness.

By contrast, conditional independence between the testimonies of the two witnesses does not hold: Knowing that one witness mentioned a specific color increases my confidence that the other witness will mention this color too. This is due to the fact that the identification of the color by the witnesses is reliable and, by consequence, it makes it more probable that the cab has in fact the identified color. However, this also increases the probability that the other witness provides the same answer since her response is influenced by the color.

Figure 4-15 depicts a possible causal structure that leads to the conditional independence between the two witnesses given the color of the cab. This common cause structure also underlies modern psychometric models (cf. Ex. 4-20, on page 136).

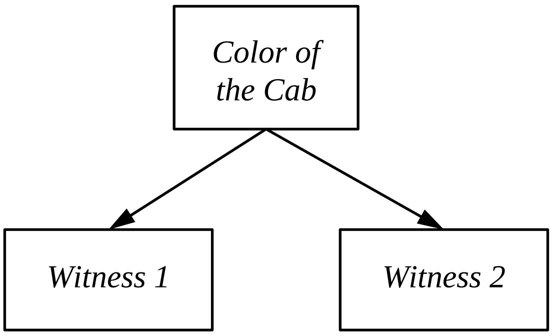


Figure 4-15: *Assumed causal structure underlying the extended cab problem that results in a conditional independence of the idenfications of the witnesses given the color of the cabs.*

(b) The different representations for representing the extended cab problem are differentially suitable for disclosing the existence of conditional independence. Contingency tables do not reveal conditional independence, whereas outcome trees with probabilities are well suited for exhibiting conditional independence.

Figure 4-16 depicts the outcome tree in probability format for the extended cab problem. Conditional independence of the testimony of the second witness from that of the first one is indicated by the fact that the conditional probabilities of identifying a blue cab as »blue« and »green«, respectively, is the same independently of whether the testimony of the first witness was »blue« or »green«. The same is true for the green cabs (cf.the numbers in red color, in Figure 4-16). Thus, the probability of the second witness responding »blue« or »green« depends on the color of the cab but not on the testimony of the other witness.

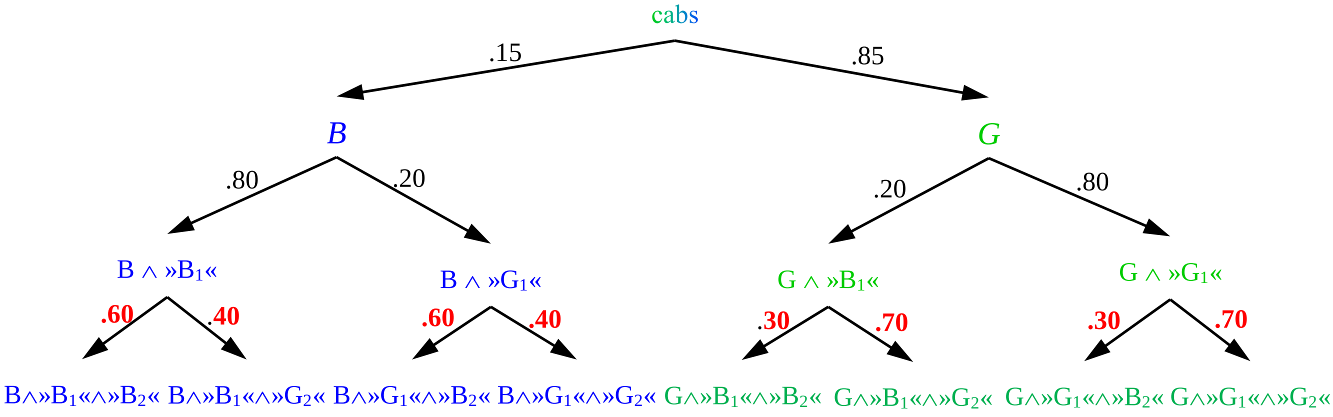


Figure 4-16: *Representation of the extended cab problem by means of an outcome tree in probability format.*

To illustrate a situation where conditional independence of the testimonies of the witnesses, given the color of the cab, does not hold, assume that the second witness knows the answer of the first witness resulting in a tendency to match her testimony with that of the other witness. In this case, the testimony of the first witness exerts a direct influence on that of the second one. In order to represent this causal scenario properly, an arrow *Witness 1* → *Witness 2* has to be added to the diagram in Figure 4-15.

For the sake of concreteness, let us modify the conditional probabilities of the outcome tree of Figure 4-16 in such a way that they reflect the tendency of the second witness to adjust her testimony to that of the first one. Figure 4-17 exhibits the resulting outcome tree.

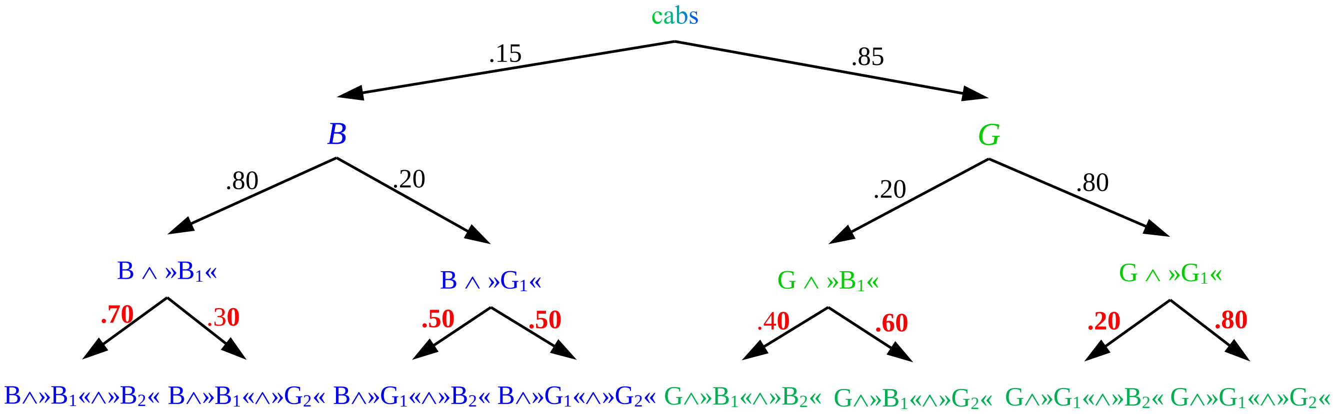


Figure 4-17: *Representation of the extended cab problem by means of an outcome tree in probability format. Conditional independence of the testimonies of the witnesses given the color of the cabs does not hold.*

Figure 4-17 exhibits the resulting outcome tree. In contrast to the outcome tree representing conditional independence, the conditional probabilities of the testimonies of the second witness are no longer the same regarless of the responses of the first witness. Rather, the probabilities reflect the tendency of the second witness to fit her response to that of the first witness. For example, in case of a blue cab, the probability of the second witness to identify the cab as »blue« is higher (.70) if the answer of the first witness was »blue« compared to the case when the answer was »green« (.50). The same holds for the green cab.

4.4.2.4 Probabilistic Operations and the Problem Space Metaphor

Newell & Simon (1972) conceptualized problem solving activity as a search through a problem space with a path connecting the initial state, given by the problem description, with the goal state, the solution. The problem space comprises four elements (Bassok & Novick, 2012):

1. A set of knowledge states: These include the initial and the goal state as well as possible intermediate states on the path connecting the initial state with the goal state.
2. A set of cognitive operators that enable the movement between states.
3. A set of constraints shaping the search through the problem space.
4. Local information associated with states about the further path one has to take through the problem space.

In case of probabilistic reasoning, the problem space may be conceptualizes by the different types of probability information. Specifically, the initial state consists of the given probabilities. In case of Bayesian reasoning these comprise the prior probabilities and the likelihoods. The goal state is given by the searched for probability which is, in case of Bayesian reasoning, the posterior probability.
The operators enabling to move between the knowledge states are the elementary probabilisitic operations of combination, marginalization and conditioning. In case of Bayesian reasoning each of these operators is required in order to arrive at the goal state. Note however that some problems do not require the application of the full set of operators. One example is the total evidence design of Shafer and Tversky (1985) that requires the operations of combination and marginalization only. (cf. Appendix: Elements of probability theory).

Following to this extensive treatment of the formal structure of Bayesian reasoning problems let us now discuss psychological theories that have been put forward to explain human performance in Bayesian reasoning and Bayesian estimation problems.

4.4.3 Cognitive Mechanisms Underlying Probabilistic Reasoning

There has been an intensive debate concerning human's ability to perform Bayesian reasoning. On the one hand it was argued by Kahneman and Tversky (1982) that human beings are unable to solve Bayesian reasoning problems, like the cab problem, without tutoring. On the other hand it has been argued that people are in principle able to solve this sort of problems. The observed difficulties stem from the fact that the probabilistic information is given in the wrong format. In the following, the principle arguments of the opposite positions will be presented. We also discuss the phenomenon of conservativism in Bayesian reasoning that seems to be in opposition to the results found for the cab problem. Finally, erroneous probabilistic intuitions (heuristics) that are applied in solving probabilistic reasoning problems are discussed.

4.4.3.1 On the Controversy between Evolutionary Psychologists and Problem Theorists

A hot debate concerning humans' ability to perform Bayesian reasoning took place at the end of the 1990th between *evolutionary psychologists* and a group of psychologists who will be called *problem theorists*.

4.4.3.1.1 Arguments of the Evolutionary Camp

The evolutionary camp argues that peoples' bad performance, observed in Bayesian reasoning problems, is due to the fact that the probability information provided in these tasks does not correspond to the way how our ancient ancestors received and processed probability information.

Probabilities and proportions are relatively recent inventions and modern theories of probability emerged in the 17th century (e.g. Hacking, 2009). Our ancient ancestors were concerned with frequencies. Specifically, in the context of Bayesian reasoning so called *natural frequencies* have turned out to improve Bayesian reasoning (Gigerenzer & Hoffrage, 1995). Here is a formulation of the cab problem in terms of natural frequencies.

Ex. 4-42: The cab problem with natural frequencies

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85 out of 100 cabs in the city are green and 15 out of 100 cabs are blue.

(b) A witness identified the cab as blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and got the following results:

- ❑ 12 out of the 15 blue cabs were identified as »blue« by the witness.
- ❑ 17 out of the 85 green cabs were identified as »blue« by the witness.

How many of the one hundred cabs that were identified as »blue« by the witness are in fact blue?

__________out of 100

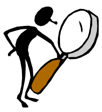
The natural frequencies in Ex. 4-42 are given by the following numbers:

- ❑ *12 out of 15* (blue cabs identified »as blue« by the witness), and
- ❑ *17 out of 85* (green cabs identified as »blue« by the witness).

Obviously these frequencies incorporate the base rates of the cabs in town. This is due to the fact that no renormalization (with respect to the basic frequency of 100) is performed. By consequence, the conditional probabilities $P\left(\text{"Blue"}\middle|\,\text{Blue}\right)$ and $P\left(\text{"Blue"}\middle|\,\text{Green}\right)$ are not expressed with respect to the basic frequencies (100) but with respect of base rates (15 and 85, respectively). This amounts to the fact that these probabilities represent the *joint probabilities*, $P\left(\text{"Blue"} \wedge \text{Blue}\right)$ and $P\left(\text{"Blue"} \wedge \text{Green}\right)$.

Obviously the cab problem with natural frequencies is much simpler than the original version of the cab problem (cf. Ex. 4-33, page 159). In fact, one can get the correct answer by dividing one of the numbers by the sum of both: $12/\left(12+17\right)$.

Natural frequencies result from a process called natural sampling.



***Concept 4-11:** Natural sampling:*

Natural sampling consists in the sequential sampling from embedded (nested) sets (or sub-populations) with the sub-populations being the result of previous sampling processes.

The absolute frequencies of different values resulting from natural sampling incorporate the base rates of the different sub-populations.

With respect to the cab problem one arrives at the natural frequencies by first sampling blue and green cabs out of the whole population of cabs, resulting in 15 blue and 85 green cabs

In the second step, the sampling is performed separately on the blue and green cabs, respectively. Out of the 15 blue cabs, 12 are sampled (=80%), and out of the 85 green cabs 17 are sampled (=20%).

Thus, the natural frequencies are the result of a sequential sampling process, with sampling being performed for different sub-populations resulting from previous sampling (cf. the sampling version of the cab problem in Figure 4-11, on page 166).

The proponents of the evolutionary argument claim that people are in fact Bayesian reasoners. However, our ancestors encountered probabilistic information in terms of (natural) frequencies only. Thus our Bayesian reasoning module fails with probabilities as input (see, e.g. Brase, 2008; Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995).

A second example of an evolutionary argument explaining why people fail for specific Bayesian problems is due to Brase, Cosmides and Tooby (1998). They consider the problem of the three cards (cf. Exercise 4-24):

Ex. 4-43: The problem of the three cards:

Given: Three cards:

1. A card with both sides being white.
2. A card with one side being white and the other one being black.
3. A card with both sides being black.

The cards are put into a box and mixed up. Then, one card is chosen at random and put on the table with one side up (The side is also chosen randomly). The side shown is white.

What is the probability that the other side is white, too?

The typical (wrong) answer to this problem is 1/2 (the correct answer is 2/3; cf. Exercise 4-24).

Brase et al. (1998) note correctly that people do not take into account that the chance of selecting a white side in case of a card with two white sides is double that of a card with one white and one black side. They argue that our reasoning mechanism has been tuned to whole objects and not to an arbitrary parsing of these objects. This explanation has been called the *individuation hypothesis*. In case of the problem of the three cards this means that people ignore the single sides of the cards since these are arbitrary parts of whole objects.

To underpin the individuation hypothesis Brase et al. (1998) performed a number of experiments in which they compare peoples' performance with whole objects to that with an arbitrary parsing of the objects. The following example demonstrates the basic logic of the experiments.

Ex. 4-44: Probabilistic reasoning with whole objects vs. an arbitrary parsing of objects (Brase, et al., 1998):

Participants are presented one of two versions of the same problem with either an arbitrary parsing version or a whole object version:

Arbitrary parsing version:

At the grocery store there are jars with straight candy canes:

- ❑ ***Jar 1*** *contains only pink candy canes. These are peppermint.*
- ❑ ***Jar 2*** *contains only yellow candy canes. These are lemon.*
- ❑ ***Jar 3*** *contains candy canes that change color and flavor in the middle.*

You take the same number of candy canes from each jar and put it in your bag.

At home your friend closes her eyes pulls a candy can out of your bag and tastes one end of it. The tasted end is peppermint.

What is the probability that the other end tastes also peppermint? ______ (correct answer: 2/3).

Whole object version:

At the grocery store there are jars with straight candy canes:

- ❑ ***Jar 1*** *contains only pink candy canes. These are peppermint.*
- ❑ ***Jar 2*** *contains only yellow candy canes. These are lemon.*
- ❑ ***Jar 3*** *contains yellow and pink candy canes in equal proportion.*

You take the same number of candy canes from each jar and put it in your bag (with equal proportion of yellow and pink canes from Jar 3).

At home your friend closes her eyes pulls a candy can out of your bag and tastes it. It tastes peppermint.

What is the probability that the candy can is from Jar 1 (and not from Jar 3)? ______ (correct answer: 2/3).

Result and Interpretation:

As one might expect the participants who received the whole object version performed better than the group working with the arbitrary parsing version (This is intuitively clear since obviously there are twice as many peppermint candy cans from Jar 1 than from Jar 3).

The results of the experiments thus confirm the individuation hypothesis.

In conclusion, evolutionary theorists argue that peoples' fallacies in Bayesian reasoning are due to the fact that the form of information in typical Baysian reasoning problems does not correspond to the way how our ancient ancestors experienced information about uncertain events. If the probability information is provided in the proper format people are able to perform Bayesian reasoning.

Let us now have a look on the arguments of the problem theorists.

4.4.3.1.2 Arguments of the Problem Theorists: Explanations Based on the Problem Representation

The proponents of this explanation (see e.g. Barbey & Sloman; 2007; Evans, 1996; Girotto & Gonzales, 2001, Sirota, Kostovičová, & Vallée-Tourangeau, 2015a,b; Sloman, Over, Slovak, & Stibel, 2003) conceive Bayesian reasoning problems as being similar to classical problems, like the Tower of Hanoi or the Water Jug Problem. According to this view the following principle is important:

> S*uccessful problem solving consists in the construction of a suitable mental representation of the problem.*

According to this view, natural sampling, as well as other methods that improve Bayesian reasoning, leads to a simplification of the problem structure. Specifically, the simplification consists in the fact that *these methods make the nested sets involved in the problem as well as the associated probabilities salient*.
To illustrate this with the natural sampling formulation of the cab problem: The problem specifies the frequencies of different sets:

1. The frequencies of the cabs of different colors in town:
   - ❑ *15 out of 100* cabs are blue, and
   - ❑ *85 out of 100* cabs are green
2. The frequencies of the embedded sets of green and blue cabs that have been identified as »blue« by the witness:
   - ❑ *12 out of 15* blue cabs identified as »blue« by the witness,
   - ❑ *17 out of 85* green cabs identified as »blue« by the witness.

Thus, the natural frequency version (as well as the outcome tree representation of the cab problem in frequency format) makes the nesting of the embedded sets and their relative size salient. This simplified representation of the problem results in higher problem solving success rates.
The problem space representation of probabilistic reasoning, discussed in Section 4.4.2.4, provides the most complete explanation:

> *Natural frequencies (and other methods) provide participants with the relevant joint probabilities. By consequence, the operation of the combination of probabilistic information need not be performed thus simplifying the problem and reducing the number of mental operations required for arriving at a solution.*

In fact, the cab problem with natural frequencies presents the relevant *joint frequencies* (namely 12 and 17). By consequence, the required conditional probability is given by dividing the relevant joint frequency (12) by the sum of both given frequencies (12+17). Thus only the operation of marginalization (summing the two frequencies) and conditioning (dividing the relevant joint probability by the marginal one) are required.

In addition, the operation of marginalization (summing up the relevant frequencies or probabuilities) is simplified since only the relevant joint frequencies are presented (and not the other two: 3 and 68).
According to this analysis the presentation of Bayesian problems using natural frequencies confounds two things:

1. The usage of (non-normalized) frequencies, and
2. The simplification of the problem by presenting the relevant joint frequencies that spares the combination of base rates and conditional probabilities and simplifies the operation of marginalization.

The two confounded factors can be separated by specifying the problem in terms of joint probabilities instead of natural frequencies.

- 12% of the cabs passing are blue and are identified as blue by the witness.
- 17% of the cabs passing are green and are identified as blue by the witness.

If the main reason for the simplification of the problem with natural frequencies is due to the usage of joint probabilities and the resulting simplification of the sequence of required operations then the specification of the cab problems by means of joint probabilities should result in similar improvements as observed for the cab problem with natural frequencies. Keren und Lewis (1999) demonstrated that this is in fact the case.
It has been argued that reasoning with natural frequencies is not Bayesian reasoning at all that (Howson & Urbach, 2006). The fact that with natural frequencies no combination of probabilistic information is required justifies this assessment (for a different view, see Brase, 2002).

4.4.3.1.3 Appraisal of the Arguments of the Two Camps

The problem space explanation provides a neat explanation of why Bayesian reasoning can be improved by different methods. Thus no unproved ad-hoc explanations and assumptions, respectively, offered by the evolutionary camp are required.
One such unproved assumption is the *individuation hypothesis* of Brase et al. (1998), discussed above, according to which our Bayesian reasoning module works properly with whole objects only. The problem of this explanation is exhibited by the following example.

Ex. 4-45: The problem of the three boxes:

Given: Three boxes:

1. Box 1 contains 2 white balls.
2. Box 2 contains 2 black balls.
3. Box 3 contains 1 white and 1 black ball.

> One of the boxes is drawn at random (each box has the same probability of being selected). From the selected box a ball is drawn at random. It is white.
>
> *What is the probability that the other ball is white, too?*
>
> This problem results in the same typical (wrong) answer as the problems with the three cards. However, in this problem whole objects are involved only and no arbitrary parsing of objects.

So why does the whole object version of Brase et al. (2008) result in an improvement? The answer to this question is obvious. The descript-ion of the problem in whole object format (cf. Ex. 4-44) makes it clear that in the bag there are twice as many peppermint sticks coming from the first Jar (The peppermint jar) than peppermint sticks coming from Jar 3 (The mixed jar, containing peppermint and lemon sticks). Thus the odds of peppermint sticks coming from Jar 1 to those coming from Jar 3 are two to one. For example, if the person takes 20 sticks from each jar, then there are 20 peppermint sticks from Jar 1 in the bag and 10 peppermint sticks from Jar 3 (the other 10 sticks from Jar 3 are lemon sticks).

The whole object version makes the fact that there are twice as many peppermint cans from Jar 1 compared to Jar 3 more salient than the arbitrary parsing version. No complicated explanation, like the individuation hypothesis, is required.

The nested set argument also applies to the Linda problem and the observed conjunction fallacy: A proper representation of the problem avoiding the conjunction fallacy has to represent the relationship between different sets, specifically the relation between *the set of bank tellers* and *the set of bank tellers being active in the feminist movement*. Clearly, the conjunction fallacy indicates that people do not represent the relation between these two sets.

Two important implications result from the conceptualization of probabilistic reasoning tasks as a sort of textbook problems:

1. Any hint leading to a more adequate representation of the problem structure should result in an improved performance.
2. The better performance with frequency formats (instead of probabilities) as well as with manipulations increasing the saliency of the probabilistic nature of the problems (cf. Section 4.6.1 and **Error! Reference source not found.**) should be due to the fact that these methods result in an improved understanding of the problem structure.

In order to test these implications Sloman et al. (2003) conducted a number of experiments. In one of their experimental conditions the relevant set involved as well as their nesting was made salient. This re-

sulted in a clear improvement independently of whether the problem was presented in frequentist of probabilistic format.
By contrast, obscuring the nesting of the sets (for example by inserting further questions between the two relevant questions for the Linda problem) led to an impairment of the reasoning independently of whether probabilities or frequencies were used.

On the basis of the foregoing discussion, the following conclusion may be drawn:

> *The main reason of the shortcomings in probability judgments is* ***not*** *due to the fact that our cognitive module that is responsible for performing inferences with uncertain information works only with frequencies, as claimed by Brase, Cosmides, Gigerenzer, Tooby, and others. Rather, the main reason consists in the fact that people are unable to construct an adequate representation of the problem structure.*
> *Any method resulting in a better representation of the problem structure, like outcome trees, contingency tables of other graphical displays discussed below, will result in an improvement of probabilistic reasoning.*

There has been a further discussion that concerns the usage of base rates in Bayesian reasoning.

4.4.3.2 CONSERVATISM VERSUS BAYES RATE NEGLECT

As discussed above, most particapants trying to solve the cab problem provided the answer 0.80, i.e., the sensitivity of the witness to identify the correct color. This indicates that people ignore the base rate information. By contrast, Edwards (1968) found that people do not use the diagnostic information sufficiently in revising the prior probabilities. He termed this phenomenon *conservativsm*.

Ex. 4-46: Conservatism in Bayesian reasoning (Edwards, 1968, 1982)

Given:

- ❑ Two bags containing 1000 chips, each:
  Bag A: 700 red and 300 green chips;
  Bag B: 300 red and 700 green chips.
- ❑ A fair coin is flipped to select one of the bags. Thus the prior probability that the selected bag is *A* or *B* is 0.5.
- ❑ From the selected bag 12 chips are drawn at random with replacement: 8 are red and 4 are green.

Required:

An estimate of the probability that Bag *A* was selected.
The typical answer is in the range from 0.7 to 0.8.

Bayesian solution:

Prior odds are 1.0 (and can thus be ignored).

The likelihood ratio is given by the binomial probabilities with parameters 0.7 and 0.3:

$$LR = \frac{(0.7)^8 \cdot (0.3)^4}{(0.7)^4 \cdot (0.3)^8} = \frac{7^4}{3^4}$$

The required probability is:

$$P\left(\text{Bag } A \middle| 8 \text{ red out of } 12\right) = \frac{7^4 / 3^4}{1 + 7^4 / 3^4} = \frac{7^4}{7^4 + 3^4} = \underline{\underline{0.967}}$$

Interpretation:
People are *conservative in adjusting their probabilities*, i.e., the adjustments are insufficient: Instead of revising the prior probability of 0.5 to about 0.97 their estimates of the posterior probabilities (between 0.8 and 0.8) are closer to the prior probabilities.

Comment on the binomial probabilities involved:
The (binomial) probabilities of getting 8 red and 4 green chips in 12 independent draws with replacement are:

$$P\left(8 \text{ red out of } 12 \middle| \text{Bag } A\right) = \binom{12}{8} \cdot (0.7)^8 \cdot (0.3)^4$$

$$P\left(8 \text{ red out of } 12 \middle| \text{Bag } B\right) = \binom{12}{8} \cdot (0.7)^4 \cdot (0.3)^8 .$$

Since the binomial coefficient $\binom{12}{8}$ is identical it cancels in computing the likelihood ratio.

The results of Tversky and Kahneman (1982) and Edwards (1968) concerning the influence of base rate information seem to contradict each other since for the cab problem base rate information is predominantly ignored, whereas in the study of Edwards the prior probabilities exert an influence too strong on participants' final estimates. This appearent contradiction may be dissolved by considering the following two arguments:

1. The cab problem is usually not conceptualized as a Bayesian updating problem by the participants. By contrast the problem given in Ex. 4-46 is stated as an updating problem.
2. In Edward's (1968) problem the likelihoods are quite complicated. It is highly questionable that untrained people dispose of a good understanding of binomial probabilities. Consequently, it is hardly astonishing that participants did not use them properly. In contrast, the likelihoods in the cab problem are much easier to understand.

This difference might explain the differential use of base rate information in the two studies.

4.4.3.3 The Monty Hall Problem and the Use of Erroneous Probabilistic Intuitions

Shimojo and Ichikawa (1989) revealed that participants solving a specific Bayesian problem make use of a number of erroneous probabilisitic intuitions (see also Falk, 1992).

One such erroneous intuition that is frequently observed in Bayesian diagnostic problems (cf. the problems of Exercise 4-12 and Exercise 4-16) consists in the confusion of the *sensitivity of a diagnostic procedure*, i.e. the probability of a positive outcome given the disease, with the *posterior predictive probability* of the disease given the outcome of the diagnostic procedure.

This error has also been observed for the cab problem that can be regarded as a Bayesian problem of the diagnostic type. In this case, the colors of the cabs are the »diseases«, and the reliability of the witness to identify the correct color of the cabs representss the sensitivity of the diagnostic procedure.

The erroneous probabilistic intuitions observed by Shimojo and Ichikawa (1989) were found in the context of a different type of Bayesian problem, *the problem of the three prisoners* (cf. Mosteller, 1965). In the following the fallacious intuitions will be demonstrated in the context of the famous Mount Hall problem that is formally identical to the problem of the three prisonsers.

4.4.3.3.1 The Monty Hall Problem

In the nineties of the last century the Mony Hall problem caused a sensation in various countries. The reason for the great publicity consisted in the fact that the correct solution ran counter to the intuition of most people, among them the famous mathematician (e.g. Paul Erdös who has got the most mathematical publications of the 20th century). This elicited, in part, highly emotional reactions.

Today, the problem is presented in most High Schools and thus educated people usually know the problem and its solution. However, in most cases a thorough understanding is missing as is evidenced by the fact that these people exhibit great difficulty to solve structurally equivalent of similar problems (cf. Exercise 4-24, and Exercise 4-25). Here is a version of the problem.

Ex. 4-47: The Monty Hall Problem:

Suppose you're on a game show, and you're given the choice of three doors (A, B, or C): Behind one door is a car; behind the others, goats. You pick a door, say Door A, and the host, who knows what's behind the doors, opens another door, say Door C, which has a goat. He then says to you, »Do you want to pick Door B?« Is it to your advantage to switch your choice?

> *Comment:*
>
> The formulation of the problem does not determine a unique probability model. Some further specifications are required (cf. the discussion below, as well as Exercise 4-19).
>
> *Intuitive solution:*
> Assume that you choose Door A (due to symmetry, it does not matter which door you choose). In this case your probability of winning the car is 1/3 (assuming that the car has the same probability of being located behind each of the three doors).
> By consequence, the probability that the car is behind one of the other two doors is 2/3.
> Now the host removes the uncertainty with respect to the question behind which of the other two doors the car is located (since, due to the constraints given by the situation of a game show, he can only open the door with a goat). Consequently, if you switch to the door that was not opened by the host you have a probability of 2/3 to win the car.

The problem can be solved easily. Let us represent the problem by means of a contingency table containing the joint (and marginal) probabilities of the two relevant random variables:

(a) The door behind which the car is located: A, B, or C.
(b) The door opened by the host: »A«, »B«, or »C«.

It is assumed that the candidate has chosen Door A (This assumption provides no restriction since the problem is completely symmetric with respect to the door chosen). In addition, it is assumed that the host has no preference for Door B or C in case of the car being located bind Door A, i.e., in this case, he chooses Door B and C with equal probability (For a representation of the problem by means of an outcome tree as well as a relaxation of the assumption of equal preference of the doors, see Exercise 4-19) [*Comment:* Clearly, the representation also assumes that each door has the same probability of hiding the car].
Tab. 4-14 reveals the joint probabilities and marginal probabilities. The table can be constructed as described in Section 4.4.1.4: We start with the marginal probabilities of the location of the car. These are found in the right-most column. Then the probabilities are »split« according to the conditional probability. For example, if the car is located behind Door A the host opens Door B and C with equal probability. Thus, $P\left(»B«\middle| A\right) = P\left(»C«\middle| A\right) = 1/2$, and, by consequence, the resulting joint probabilities are 1/6 each. In case of the car being located behind Door B, the host is forced to open Door C. The conditional probabilities are thus $P\left(»B«\middle| B\right) = 0$ and $P\left(»B«\middle| C\right) = 1$. An

analogous reasoning applies to the case where car is located behind Door C. Note also that the logic of the game excludes the possibility that the host opens the chosen Door A.

Tab. 4-14: *Joint probabilities of the location of the car and the door opened by the host for the Month Hall problem under the assumption that the candidate has chosen Door A and the host has no preference for one of the doors (B or C) in case of the car being located behind Door A.*

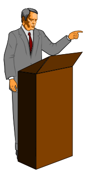
	Door opened by the host			
Door with car	»*A*«	»*B*«	»*C*«	Σ
A	0	1/6	1/6	1/3
B	0	0	1/3	1/3
C	0	1/3	0	1/3
Σ	0	1/2	1/2	

On the basis of the joint probabilities the relevant conditional probabilities are easily computed:

$$P(A\,|\,»B«)=\frac{P(A,»B«)}{P(»B«)}=P(A\,|\,»C«)=\frac{P(A,»C«)}{P(»C«)}=\frac{1/6}{1/2}=\frac{1}{3}\text{, and}$$

$$P(B\,|\,»C«)=\frac{P(B,»C«)}{P(»C«)}=P(C\,|\,»B«)=\frac{P(C,»B«)}{P(»B«)}=\frac{1/3}{1/2}=\frac{2}{3}\,.$$

Thus the probability of getting the car if the candidate stays at Door A is only half the probability of winning the car in case of shifting to the other door (*B* or *C*).



Comment 4-8: Possible objections

Nickerson (1996) and Gigerenzer (2001) claim that the problem has not been specified sufficiently. Specifically, Gigerenzer (2001) put forward the following argument:

»… the correct answer – stay or switch – depends on the knowledge one has of the situation and the assumptions one has to make in the absence of full knowledge. As before, these assumptions were not specified in the original problem version. For instance, it is essential to know whether or not Monty always offers a candidate the possibility of switching doors. If he were to offer switching only when the candidate was standing in front of the winning door, the normative strategy would be not to switch.« (Gigerenzer, 2001, pp. 99-100).

This type of argument misses one important point: The problem is meant as a mathematical brainteaser, and the fact that the host might behave unfair has nothing to do with the mathematics. Thus Gigerenzer's speculation intentionally misunderstands the intent of the inventor of the problem by drawing attention on non-mathematical aspects thus making the problem completely trivial.

Gigerenzer's argument also runs counter to the conversational principles of Paul Grice that Gigerenzer never gets tired to cite.

According to these principles people communicating with one another follow a number of maxims that enable a convenient way to exchange ideas.

According to one of these principles, the *maxim of quantity*, everything that is relevant for the understanding of the problem structure has to be mentioned. Thus, if the host pursuits an unfair strategy by offering a choice only in case of the car being located behind the door chosen by the candidate this has to be mentioned since otherwise the situation could not be understood properly. However, in this case, the whole problem would lose its status as problem at all.

It also important to note that Gigerenzer's argument cannot explain the difficulty of structurally similar problems that are not subjected to Gigerenzer's far-fetched speculation.

Nickerson (1996, pp. 418-419) also claims that the host might choose randomly between the doors not selected by the candidate thus permitting the possibility that the opened door reveals the car. According to the present view, this claim is far-fetched too, and the logic of the game excludes this behavior of the host.

A complete specification of the problem includes additional information about:

1. The uniform prior distribution of the doors behind the car and goats are hidden;
2. That the host must open one of the two door not chosen by the contestant containing a goat;
3. The fact that in case of the car being hidden behind the door chosen by the contestant, the selects one the other two doors with equal probability.

Presenting the complete specification of the problem does not result in any improvement of probabilistic reasoning (Krauss & Wang 2002).

Let us now turn to possible fallacious intuitions that are used in solving the problem.

4.4.3.3.2 Erroneous Heuristics Applied to the Monty Hall Problem

Shimojo and Ichikawa (1989) argue that people's intuitive probabilistic reasoning is guided by erroneous intuitions about probabilities (see also Falk, 1992). Specifically, they identified the following three subjective »theorems« (or heuristics) guiding participants reasoning:

1. *Number of cases:* If the number of remaining alternatives is N, the probability of each alternative is $1/N$.
2. *Constant ratio:* If an alternative is eliminated the ratio of the probabilities of the remaining alternatives is the same as the ratio of their prior probabilities.
3. *Irrelevant therefore invariant:* If it is certain that (at least) one of several alternatives will be eliminated, and the information specifying which alternative to be eliminated is given, it does not change the probabilities of the remaining alternatives.

Let us check the predictions that are made by the three intuitive theorems in case of the Monty-Hall problem:

1. The *number of cases* heuristic will predict that the reasoner assumes that the probability of the price being behind the chosen door is 1/2 since, after opening one door there remain two alternatives.
2. The *constant ratio* heuristic makes the same prediction as the number of cases heuristic since the prior probabilities of the remaining doors are identical and, the odds of the price being behind one of the two doors are thus 1:1, resulting in a probability of 1/2.
3. The *irrelevant therefore invariant* heuristic makes the same prediction as Bayes theorem (p = 1/3) since by opening a door and thus eliminating one alternative does not provide any new information. By consequence, the probability of the price being located behand the selected door should not be changed on the basis of the given information.

Thus, the *number of cases* and the *constant ratio* heuristic lead to a wrong answer whereas the *irrelevant therefore invariant* heuristic provides the correct answer. However, the correctness of the latter depends crucially on the response strategy of Monty in case of the car being located behind the door selected by the candidate: The strategy works correctly in case of Monty opening Door *B* and *C* with equal probability. If, by contrast, Monty is biased by always opening Door *B* (or *C*), if possible then the *irrelevant therefore invariant* heuristic fails (Exercise 4-19). The same happens in case of unequal priors, i.e. in case of unequal prior probabilities concerning the location of the car. Also in this case, the posterior probability need not be the same as the prior probability (cf. Shimojo & Ichikawa, 1989, Appendix 1, for a formal analysis in case of the three prisoners problem and Falk, 1992, for the general case [see also Exercise 4-11]).

Shimojo and Ichikawa (1989) also identified *superior subjective theorems.* These have a strong intuitive appeal and may be relevant for selecting subjective theorems. As an example they present the following belief:

> *If an alternative (with a probability greater than zero) is eliminated from the set of possible alternatives the probability of the remaining alternatives can never decrease.*

This belief can result in an erroneous reasoning too, as demonstrated by a revised version of the problem of the three prisoners (with unequal prior probabilities) created by Shimojo and Ichikawa (1989).

In conclusion, the Mony-Hall problem elicits a number of erroneous intuitions that can but need not result in a wrong solution of the problem.

4.5 Dual Process Theories of Judgment and Decision Making

In recent years dual process theories (DPT) of judgment and decision making have become increasingly prominent (see, e.g., Evans, 2006, 2008; Evans & Stanovich, 2013a,b; Kahneman, 2011; Kahneman & Frederick, 2002; Sloman, 1996; Stanovich & West, 2010; Thompson, 2013). In this section, I first provide a short description of the approach together with concrete examples that demonstrate the idea (Section 4.5.1).

Dual process theories have been criticized by different scientists (Gigerenzer & Regier, 1996; Keren, 2013; Keren & Schul, 2009; Kruglanski, 2013; Kruglanski & Gigerenzer, 2011; Osman, 2004, 2013). The second section discusses critical aspects and provides an evaluateon of DPT (Section 4.5.2).

4.5.1 Dual Process Theories (DPT)

The dual process approach concretizes the distinction between *intuitive* and *rational deliberate reasoning*. It assumes the existence of two different modes of processing called *Type 1* and *Type 2*.

4.5.1.1 Characterization of Type 1 and Type 2 Processes

The two types differ with respect of to various aspects. Tab. 4-15 is a reproduction of Evans (2008, Table 2 on p. 257) that lists attributes of the two types of processing grouped around four categories.

Notation 4-7:

The labels *Type 1* and *Type 2* are used most often in recent publications. Another frequently used naming uses the terms *System 1* and *System 2* processes. This latter labeling has been criticized since it identifies the two different types of processses with two different cognitive systems, an assumption that most researchers do not subscribe.

The characterization presented in Tab. 4-15 seems to be quite evident. For example, Type 1 processes are, in general, performed fast, effort-less and without conscious knowledge, whereas those of Type 2 are slow, require effort and are consciously penetrable. However, some of the labels presented in Tab. 4-15 require some explanation:

(1) Type 1 processes are due to *implicit knowledge*, i.e., knowledge that cannot be verbalized.
(2) Type 1 processes are of *high capacity* which signifies that they are not restricted by means of capacity or attentional limitations (see also the feature *independence of working memory*).
(3) The *modularity* of Type 1 processes means that they are apt to solve domain specific information processing tasks (see also the characteristic of *domain specificity* and *contextualization*). Type 2 processes are, on the other hand, relevant for general problem solving (*fluid intelligence*).
(4) Type 1 processes are elicited *by default*. However, they may be *inhibited* by Type 2.
(5) Type 1 processes are assumed to solve practical problems that are relevant from an evolutionary perspective, for example to detect cheaters (Cosmides & Tooby, 1992). Thus, they are assumed to be *pragmatic*. By contrast Type 2 processes are subjected to logical principles.
(6) Type 1 processes are assumed to not underlie individual differences, whereas Type 2 processes differ considerably between individuals.

Tab. 4-15: *Characterization of two types of processing (according to Evans, 2008)*

Type 1	Type 2
Consciousness	
Unconscious (preconscious)	Conscious
Implicit	Explicit
Automatic	Controlled
Low effort	High effort
High capacity	Low capacity
Rapid	Slow
Default process	Inhibitory
Holistic, perceptual	Analytic reflective
Evolution	
Evolutionary old	Evolutionary new
Evolutionary rationality	Individual rationality
Shared with animals	Uniquely human
Nonverbal	Linked to language
Modular cognition	Fluid intelligence
Functional Characteristics	
Associative	Rule based
Domain specific	Domain general
Contextualized	Abstract
Pragmatic	Logical
Parallel	Sequential
Stereotypical	Egalitarian
Individual Differences	
Universal	Heritable
Independent of general intelligence	Linked to general intelligence
Independent of working memory	Limited by working memory capacity

The assumption of two types of processing may be traced back to Freud (1856-1939) and William James (1842-1910), the forefather of American psychology. Freud distinguishes between primary and secondary processes with the first type being characteristic for the *Id*, and revealing themselves, for instance, in dreams. They follow the pleasure principle. By contrast, secondary processes are characteristic for the *Ego* and conform to the the reality principle. William James proposed a dichotomy between an associative and a deliberative mode of thinking.

4.5.1.2 APPLICATION OF DPT TO EXPLAIN JUDGMENT AND REASONING ERRORS

In applying DPT to judgment and decision making an additional assumption is made, which is called the *default interventionist assumption* (*DI*). According to DI Type 1 processes are the default that are ellicited by typical judgment, reasoning, and decision problems. The result of Type 1 processes may be overridden by the invocation of Type 2 processes that that may (or may not) correct for the error.
The invocation of Type 2 processes is subject to the following requirements:

4. The result of Type 1 processing is not satisfying to the reasoning person. For example, she has the feeling that something is not correct or the something is missing (Evans & Stanovich, 2013a,b; Kahneman and Frederick, 2002). Evans (2006) talks of the *principle of satisficing* (This is due to Simon's satisficing account of bounded rationality, cf. Section 4.2). This conforms to the idea that humans are cognitive misers that try to avoid cognitive effort that is usually associated with the invocation of Type 2 processing.
5. The reasoner has the necessary mindware, i.e. the knowledge and cognitive abilities, to generate a better representation of the situation using Type 2 processing.
6. The (additional) cognitive ressources required by Type 2 processes have to be available. Due to time constraints and/or load of working memory, because of other ongoing processes requiring working memory, the capacity of working memory required by Type 2 processes may not be available within the available time. By consequence, Type 2 processing is not successful in correcting the outcome of Type 1 processing.

Let us now look at some examples:

Ex. 4-48: DPT and the conjunction fallacy:

The representativeness heuristic which is assumed to be Type 1 processing is invoked (Kahneman & Frederick, 2005). In addition, Type 2 processing that checks for set inclusion does not correct for the erroneous result stemming from Type 1 mode of processing.
Increasing the saliency of the nesting of sets involved in the Linda problem fosters the Type 2 processing, i.e. the checking of relative set sizes, and thus results in a reduction of the conjunction fallacy (Tversky & Kahnemann, 1983; Sloman, Over, Slovak, & Stibel, 2003).

Ex. 4-49: DPT and the cognitive refection test (CRT):

Frederick (2005) designed a test consisting of 3 items that he termed the Cognitive Reflection Test (CRT). Here are the three items (Frederick, 2005, p.27).

(1) *A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball.*
How much does the ball cost? ____ cents.

(2) *It takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets. ____ minutes.*

(3) *In a lake, there is a patch of lily pads. Every day the patch doubles in size. If takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?*
____ days.

The items of the CRT were intended to elicit a wrong intuitive response that spring to mind immediately. Due to the fact that people are *cognitive misers*, i.e., they avoid spending cognitive effort if not required. Consequently, they do not invoke effortful Type 2 processing to correct for the wrong answers. Thus the CRT was regarded as to test for the capability to override a predominant response by engaging into a more effortful mode of processing.
Frederick (2005) correlated the sum of the correct responses in CRT (ranging from 0 to 3) with different forms of decision behavior, like the propensity to postpone an immediate lower reward in favor of higher reward in the future or the preference for risk. He found a relationship between CRT measures and the tendency to postpone gratification for short-term choices but not for longer time horizons. In addition, people with higher CRT measures were more willing to accept risk. Frederick (2005) also found a sex difference with respect to the CRT: Men scored in generally higher than women. The reasons for this difference are unclear.

With respect to judgment and decision biases the CRT turns out to be a better predictor of judgment and decision errors than traditional tests of intelligence (Toplak, West, & Stanovich, K. E., 2011). Not only does the CRT explain more variance than traditional intelligence items but it also exhibits incremental validity: If the CRT measure is included as a predictor in a regression analysis additionally to the measures of intelligence it explains additional variance that was not explained by the measures of intelligence. This may be interpreted as an indication that cognitive biases found in traditional judgment tasks are due to Type 1 processing that is not corrected by invoking Type 2 processes.

Ex. 4-50: DPT and belief bias in syllogistic reasoning:

Consider the following syllogism from Evans, Barston, and Pollard (1983):

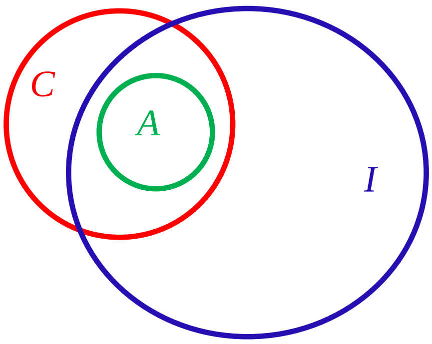
No addictive things are inexpensive.	(Premise 1)
Some cigarettes are not inexpensive.	(Premise 2)
Some addictive things are not cigarettes.	(Conclusion)

71% of the participants accepted the validiy of the conclusion which is invalid since it is possible to create a confirguration of sets that conform the two premises but not to the conclusion (cf. Figure 4-18). The existence of such a model proves that the inference is invalid. A correct conclusion was: *Same cigarettes are not addictive things.*

The high rate of erroneous judgments can be explained by means of belief bias: The conclusion is highly believable and thus the syllogism is accepted as being correct. This explanation is confirmed by the results concerning all possible combinations of the plausibility of the conclusion and the validity of the syllogism (cf. Tab. 4-16):

Tab. 4-16: *Accepatance rate of different types of syllogisms as a function of plausibility of the conclusion and the validity of the syllogism.*

Combination of plausibility and validity	Acceptance rate
Believable & valid	89%
Unbelievable & valid	56%
Believable & invalid	71%
Unbelieveable & invalid	10%



C = cigarettes
A = addictive things
I = inexpensive things

Figure 4-18: *Venn diagram representing a model that conforms to the premises but not to the conclusion of the syllogism* *of Evans et al. (1983), thus showing that the conclusion is not valid.*

The high rate of erroneous acceptance of invalid syllogisms with plausible conclusions can be explained using DPT and DI (default interventionist assumption) as follows:

The Type 1 process applied by default consists in checking the plausibility of the conclusion. Since the outcome is satisfactory Type 2 processes for correcting the result are not invoked.

In addition, people may lack the ability to perform the task required simply because they don't know how to check the validity of the conclusion of a syllogism.

These examples demonstrate that the DPT approach with DI may be applied to quite different types of reasoning and judgmental errors. However, the DPT approach has been criticized recently by various authors. Let us next turn to a critical evaluation of DPT.

4.5.2 Criticism and Evaluation of Dual Process Theories

The previous examples reveal that, on first sight, the DPT framework has some intuitive appeal. This has also been conceded by critics of the approach (cf. Keren & Schul, 2009).A closer look reveals a number of shortcomings, however. Early critics of the approach have stated that the approach is not unified. In fact there exist different versions of DPT theories assuming different characteristics of Type 1 and Type 2 processes (cf. Keren & Schul, 2009; Osman, 2004).

As a reaction to the criticisms Evans and Stanovich (2013a, b) have provided a modified version of DPT. In the following, this recent version of DPT will be discussed and evaluated.

4.5.2.1 Characterization of DPT According to Evans & Stanovich (2013)

The modification of the DPT by Evans & Stanovich concerns to aspects:

1. Instead of the many characterizing features associated with Type 1 and 2 processes (cf. Tab. 4-15, p. 214) a small set of defining features are introduced. Specifically, the idea that Type 1 processes typically result in biased results whereas Type 2 processes lead to correct results is rebuted.
2. DPT as a general framework is assumed to be a meta-thory that cannot be refuted by means of specific empirical evidence.

The first modification assumes that the following features define the two different types of Processes:

- *Type 1* processes are defined by their autonomous character and the fact that they require no working memory ressources. The autonomous character is revealed by the fact the process, once triggered by the cues of the situation executes without further control by the subject (see also Thompson, 2013).

- The defining features of *Type 2* processes are their reliance on working memory and their counterfactual nature, i.e. they are able to process models and representation that may be detached from the actual reality. Thus Type 2 processes can be used for mental simulation.

The second modification concerns the empirical testability and the evaluation of the the DPT framework. The framework is considered as a meta-theory that cannot be directly tested empirically. Only specific versions of the theory may be tested. The value of the general framework can however be evaluated according to its fruitfulness for generating new predicitons and their unifying nature of different approaches within a single framework.

4.5.2.2 Criticism of the Modified DPT Framework

The present criticism is concerned with two issues:

1. The classification of a continuum of processes into two different types of processes.
2. The problem of the explanatory power and scientific value of the framework.

Let us now investigate the two cricisms more closely.

4.5.2.2.1 Two Qualitative Different Processes versus a Single Process of Different Degree of Autmaticity

Each single critic, cited above, has put forward the criticism that DPT provides a somewhat arbitrary split of an inherent continuum. As already noted by the theories concerning the difference between automatic and controlled processes (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977) the degree of control and working memory load is a question of *quantity* and not of *quality* since intensive training results in an automatication. This is accompagnied by a lower working memory load as well as lower control (cf. Anderson, 1982, 1987). Consequently, the amount of cognitive load and control depends on the degree of overlearning.

It has also been argued that there exist processing modules for higher cognition, like a cheater detection module, that, similar to the module processing visual information, is »hard-wired« in the brain, works automatically, and is immune to correction by means of learning (Cosmides & Tooby, 1992). Processing within such a module might be classified as Type 1 processing. However, the proposition of processing modules for higher cognition was met with heavy criticism (e.g. Atran, 2001). Consequently, the splitting of a continuum of processes that are characterized by different degrees of automatization into two different types of processes is arbitrary.

Evans & Stanovich (2013a, b) also cite neurophysiological evidence in favor of the two types of processing indicating that different brain regions are acitive for the two types of processes. However, the evidence is everything else but conclusive (cf. Osman, 2004, Kruglanski, 2013). There is an additional problem: People may use different strategy to solve a problem with different strategies involving different degrees of cognitive ressources. For example, in syllogistic reasoning people may either using a plausibility check of the conclusion or mentally manipulate Venn diagrams in their head (cf. Ex. 4-50 on p.218). In the first case, people retrieve the relevant semantic knowledge that is, due to high expertise, easily accessible and thus requires little ressources, whereas in the second case people have to manipulate mental models which may, due to little training, require much more cognitive ressources. It is thus not surprising that different regaions in the brain are activated. However, the two different strategies do not map neatly onto the distinction between Type 1 and Type 2 processing since the retrieval of information from memory may also require cognitive ressources in case of processing new information about an unfamiliar topic where the relevant information does not come to mind fluently. Moreover, a person highly trained in manipulating Venn diagrams will require much less ressources for performing the task than a novice.

4.5.2.2.2 On the Explanatory Power of DPT

In order to assess the possible contribution of DPT to the explanation of cognitive biases it is useful to first consider the issue of what makes up a good explanation. A good explanation of judgment and decision biases consists of two parts:

(a) An analysis of the task: What makes up the problem and which steps have to be performed in order to solve the problem?

(b) An investigation of how the problem in question is represented by the reasoning person, and how this conceptualization of the problem and the strategies (or steps in the problem space) used for solving the problem result in the observed error or bias.

Now, in order to provide any new insight the DPT has to deliver new facettes that goe well beyond the information given by the traditional approach of analysing the problem conceptualization of the reasoning person. Apparently Evans and Stonovich (2013b) think that DPT, as a sort of meta-theory, provides a unifying framework that integrated different conceptions. Let us discuss two examples and investigate the possible contribution of DPT to our understanding of judgmental errors, additionally to traditional explanations.

First, consider the Linda problem (cf. Ex. 4-11). According to the traditional analysis, the task consists in predicting the profession and personal characteristics of Linda using the cues presented in the short description of Linda. A proper representation of the problem has to include the relationships between sets, specifically, that the set of *bank* *tellers that are active in the feminist movement* is included in the set of *bank tellers*. The presence of the conjunction fallacy indicates that participants' representations miss this important aspect. Rather they predominately base their judgment on comparing the description of Linda with the description of the different jobs assigning ranks according to how closely the two fit together due to their subjective theories (reaoning by means of represenativeness): Since a *bank teller that is active in the feminist movement* fits the description of Linda better than a *bank teller* it is assumed to be more probable. Increasing the saliency of the set relationship increases the rat of correct solutions (cf., Tversky and Kahneman, 1983; Sloman, Over, Slovak, & Stibel, 2003).

What new insights might be gained by applying the two-process framework? In order to apply the approach one has to assume that the similarity judgment (or the judgment by representativeness) is a Type 1 process whereas assessment of set inclusion has to be characterrized as a Type 2 process. However, this assignment is completely arbitrary since it is questionable whether the deliberate consideration of personal characteristics for judging the profession of Linda requires less cognitive resources than checking for set inclusion. In fact this type of explanation does not contribute anything to our understanding of the error. It rather obscures the original explanation since it is less specific and more questionable.

As a second example, consider the first problem of the CRT (cognitive reflection test). Formally the problem can be represented by a system of two equations:

$$\begin{aligned} x + y &= 1.10 \\ x - y &= 1.00 \end{aligned},$$

where *x* is the price of the bat and *y* is the price of the ball. Solving this system of equations renders the correct results: $x = 1.05$, $y = 0.05$.

People who fail to find the correct solution might construct the problem as follows: First, they represent the relationship $x + y = 1.10$, which is provided in the description. In addition, the quantity of 1.00 Dollar is mentioned. Since the splitting $x = 1.00$ and $y = 0.10$ fits nicely to what is given (and intuitively quite appealing, respectively), they come up with this erroneous solution. De Neys, Rossi, and Houdé (2013) argue that people perform a sort of substitution by replacing the critical relation *more than* (in the bat costs 1 Dollar more than the ball) by the relation with *the bat costs 1 Dollar*.

Obviously, they do not check that $1.00 - 0.10 = 0.90$ and not 1.00 Dollar. This may be due to the fact that the objective given by the task description (that the *bat has to be 1 Dollar **more** than the bat*) is simply ignored or that they do not want to spend further effort because of the intuitive appeal of the first solution that comes to mind. Thus it might be argued that the high plausibility of the first solution that comes to mind prevents a more detailed examination of the problem (satisficing principle). Interestingly, De Neys et al. (2013) showed that people are, at least partly, aware that they might not have got the answer correctly since compared to the simpler problem (note that no difference is involved):

> *A banana and a magazine together cost $2.90. The magazine costs $2.00. How much does the banana cost?*

people were less confident that they had correctly solved the original ball and bat problem. Thus, they had a feeling that their structuring of the original problem might not have been correct. This indicates that people do not simply substitute unconsciously the more complex problem by the simpler one that assumes *that the bat costs $1.00*. However, the result raises the question why people do not spend more effort to solve the problem correctly if they are uncertain about their solution: Are they unable to provide a better representation of the problem (assuming that further effort will not result in a better solution) or are they simple reluctant to investigate more effort to develop a more precise representation of the problem structure? Sinayev and Peters (2015) demonstrate that success in the CRT is linked to the numerical abilities of the participant which seems to be a key factor. It may be concluded that people with greater numerical ability are able to generate a better representation of the problem that helps them to avoid the errors.

Similar to the Linda problem DPT does not contribute any additional insight. It is unreasonable that the processes resulting in an erroneous solution are automatic and require no or less cognitive resources than checking whether the difference between assumed prices for the bat and the ball is really 1.00 Dollar. In addition a person having received some training in solving a system of linear equations might have found the solution with little effort. Kruglanski (2013) has thus proposed a default interventionist approach based on a single process where the knowledge and expertise of the reasoning person is the relevant factor that determines the strategy used.

In summary, the two examples presented do not reveal any significant contribution of DPT to our understanding of cognitive biases and judgmental errors. The traditional approach of analyzing the representation of the problem by the reasoning person as well as of the mechanisms she uses to solve the problem provides a more versatile method than the reference to a highly questionable dichotomy of two qualitative different processes.

4.6 Improving Probability Judgments

A number of attempts have been undertaken to improve probability judgments and to reduce biases in probabilistic judgments. We have already encountered one case of the improvement of probability judg-

ments in the context of base rate neglect: The usage of *causal base rates* resulted in an increased influence of base rates on participants' judgments (cf. Ex. 4-29, on p.151). This works also for the cab problem: Replacing the information that 85% of the cabs being green and 15% being yellow by the causally relevant information that *the green cabs are involved in 85% of the accidents in town whereas the blue ones are entangled in 15% only*, results in an increased impact of base rates thus reducing the bias of the probabilistic judgment.
In this section we investigate various methods for improving probability judgments. The beneficial effect of these methods can be explained by their influence on the *representation of the problem situation*. For example, in the case of causal base rates, people realize that base rates convey important information, and, consequently, must not be ignored. Thus the base rates are incorporated into the representation of the problem situation. The following methods that are related to improved probabilistic judgments and reasoning, are discussed:

1. Methods that increase the saliency of the random aspect of the problem (Section 4.6.1),
2. Graphical representations (Section 4.6.2)
3. Formal training (Section 4.6.3), and
4. Open-mindedness, intelligence and critical thinking (Section 4.6.4).

Whereas the first three points are solely concerned with methods of training the final one considers personality traits.

4.6.1 Increasing the Saliency of the Random Aspect of the Problem

One criticism of the heuristics and biases approach states that the problem formulations of Tversky and Kahneman represent the random aspect of the problem insufficently. By consequence, increasing the saliency of the random nature of the events involved should result in a reduction of the biases observed.
One possibility of improving probabilistic judgments consists in the employing frequencies instead of probabilities. The usage of frequencies instead of probabilities has resulted in improved probability judgments in various domains:

- ❑ Using frequencies leads to a reduction of the conjunction fallacy (Tversky & Kahneman, 1983; Hertwig & Gigerenzer, 1999; Sloman, Over, Slovak, & Stibel, 2003).
- ❑ Using relative frequencies instead of probabilities also improved Bayesian reasoning (Cosmides & Tooby, 1996).
  However, the results of this study could not be replicated. Specifically, whereas in the study of Cosmides and Tooby (1996) the rate of correct solutions was over 70% Sloman et al. (2003) found only between 30% and 50% solutions for exactly the same problem formulations with a similar sample of students.

In addition, Girotto and Gonzales (2001) demonstrated that for one of the problems an answer quite close to the correct solution may be obtained by means of a short-cut that is unrelated to Bayesian reasoning.

- Training of Bayesian reasoning using frequencies instead of probabilities results in higher solutions even after weeks (Sedlmeier & Gigerenzer, 2001) as well as in better transfer performance (Hoffrage, et al., 2015).

However, the usage of frequencies can also result in lower rates of solutions for Bayesian problems. Specifically, if the sample sizes in the problem formulation and sample size of the final question are not the same the number of correct solutions decreases (Ayal & Beyth-Marom, 2014).

Another method to enhance the perception of the random aspect consists in having participants perform random sampling and random simulation. For example, Gigerenzer, Hell, und Blank (1988) demonstrated that the base rate neglect vanishes if participants are allowed to draw the sample themselves. The usage of computer programs for simulating the process of random sampling becomes more and more common in the context of teaching probability in schools (cf. various articles in Chernoff & Sriraman, 2014).

However, in most cases it is not sufficient to simply state in the problem description that a process of random sampling is involved. In general, this does not result in an improved reasoning. Thus, the verbal highlighting of the random nature of the events involved is not sufficient for improving performance.

In general, it is a useful strategy to approach a problem involving uncertain events from a sampling perspective, i.e. to imagine how the results may be obtained by means of sampling. This approach requires the realization of all the actions that have to be performed as well as of the probabilistic aspects in terms of frequencies. Consequently, the basic structure of the probability model underlying the problem as well as the underlying assumptions become obvious (cf. the sampling version of the cab problem in Section 4.4.1.3).

4.6.2 Graphical Representations

As one might expect, graphical representations that exhibit the sizes of different sets and relations between sets can be of great help in improving probabilistic judgments. Two types of graphical representations have already been presented: Venn diagram (see, for example, Figure 4-4, on p. 132, and Figure 4-9 on p. 162) and outcome trees (see, for example, Figure 4-10, on p. 163, and Figure 4-16, on p. 196).

The significance of graphical representations for improving probabilistic reasoning has been investigated predominantely in the context of Bayesian inference. Before we turn to the discussion of these studies two examples are presented. They illustrate the beneficial effect of using graphical displays in risk communication and in avoiding base rate neglect.

	Ex. 4-51: Communication of medical risks
	Aspirin has beneficial effect with respect to the protection against heart attacks and strikes. Specifically it reduces the risk of getting a heart attack or stroke by about 13%.
	The graphical display of Figure 4-19 gives an impression of how big the effect really is: About 8 out of 100 people that do not take aspirin get a heart attack or stroke. For the group with aspririn the number is about 7. Thus the relative reduction is $(8-7)/8=0.125$ or 12.5%.
	The graphical display elucidates that the beneficial effect of taking aspirin might be lower than the number signifies.

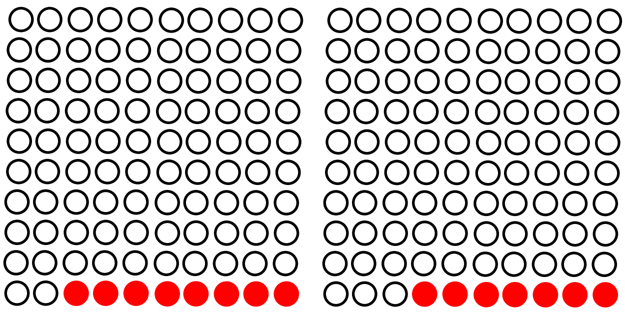


Figure 4-19: *Graphical representation of the effect of aspirin on heart attacks and strokes (from Meder & Gigerenzer, 2014, page 140).*

The next example illustrates the use of a graphical representation in case of approximate reasoning, i.e. the exact numbers are not available.

	Ex. 4-52: A base rate problem: The shy student
	You meet a student of whom you know that he studies either economics or he is a phil. diss. student in mathematics.
	Later on you receive the further information that the student is introverted.
	What do you think: Is the guy a student of economics or rather a phil. diss. student in mathematics?

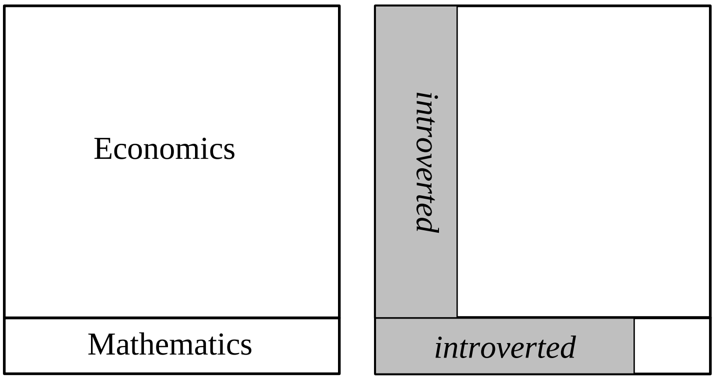


Figure 4-20: *Graphical representation of the problem of the shy student. The sizees of the rectangles represent the sizes of the relevant sets.*

Figure 4-20 represents the situation. It makes clear that, despite the fact that students in mathematics have a higher proportion of introverts than students in economics, the probability that the student you met studies economics is higher than the probability that the student you met is a phil. diss. student in mathematics. The reason is that there are much more students in economics than phil. diss. students in mathematics.

Let us now take a short look on graphical aids that have been used in various studies for improving Bayesian inference. In addition to using trees and contingency tables frequency grids and roulette wheels have been used for representing the different sets as well as their size.

Sedlmeier and Gigerenzer (2001) instructed their naïve participants to translate the given probabilities to frequencies and to use a frequency grid for representing the relevant frequencies. The frequency grid represents the different sets as well as their frequencies.

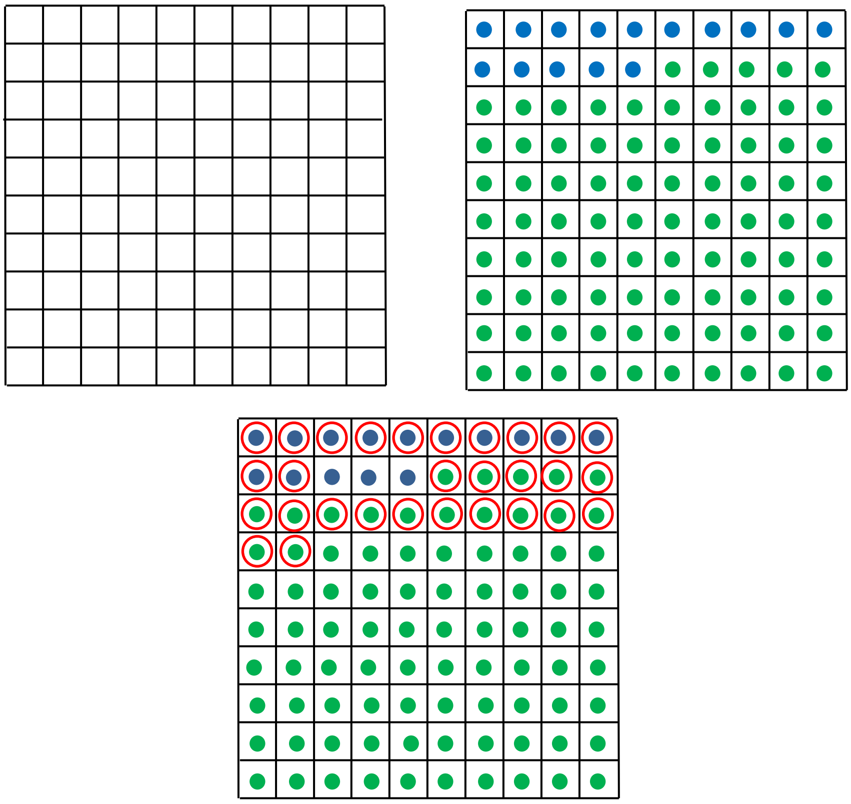


Figure 4-21: Frequency grids for representing the different sub-populations for the cab problem.

Error! Reference source not found. depicts various stages in filling a f requency grid that can be used subsequently to compute the relevant conditional probability: The upper left grid depicts the empty grid consisting of 10×10 cells. The upper right grid shows the frequency grid filled with green and blue dots that represent the green and blue taxis. In the lower grid the red circles represent the 12 blue and 17 blue cabs that have been identified as blue by the witness.

The construction of the configuration mirrors the process of natural sampling: First, the two sub-populations are constructed (upper right grid). Second, for each of these two sub-populations the embedded sub-populations of cabs that have been recognized as blue by the witness are identified (lower grid). By consequence, the final setup represents the *joint frequencies* with respect to the whole population (cf. Section 4.4.3). Thus, the relevant posterior probability can be computed by means of a simple calculation: Divide the number of blue cells with a red circle by the number of cells with a red circle.

The frequency representation turned out to be quite favorable with respect to the long term stability of the training: After five weeks participants trained to use frequency grids for solving Bayesian problems similar to the cab problem were 100% correct. The method clearly outperforms outcome trees representing relevant probabilities (cf. Figure

4-10 on page 163). It performs about equally well as trees representing the natural sampling and natural frequencies, respectively (cf. Section 4.4.3).
A different graphical representation was used by Yamagishi (2003). In this case, a roulette-wheel representing the relevant joint probabilities was used (cf. Tab. 4-17). Yamagishi (2003) investigated how different aids improve Bayesian reasoning in problems that are structurally equivalent to the Monty-Hall and Three-Prisoners problem, respectively (cf. Section 4.4.3.3.1)

Ex. 4-53: Gemstone problem and roulette-wheel representation (Yamagishi, 2003)

The following problem (with frequencies or probabilities) was presented to participants in Experiment 1:

Frequency version:

A factory manufactures 1200 artificial gemstones daily. Among the 1200, 400 gemstones are blurred, 400 are cracked and 400 contain neither. An inspection machine removes all cracked gemstones, and retains all clear gemstones. However, the machine removes half of the blurred gemstones.
How many gemstones pass the inspection and how many among them are blurred?

Probability version:

A factory manufactures artificial gemstones. Each gemstone has a 1/3 chance that it is blurred, 1/3 chance that it is cracked, and 1/3 chance that it contains neither. An inspection machine removes all cracked gemstones, and retains all clear gemstones. However, the machine removes 1/2 of the blurred gemstones.
What is the chance that a gemstone is blurred after inspection?

The problems were presented either with or without a roulette-wheel. The inner circle of the roulette-wheel is partitioned into three sections of equal size and represents the output of the production, specifically that an equal proportion of clear, blurred and cracked gemstone are produced (cf. Tab. 4-17).

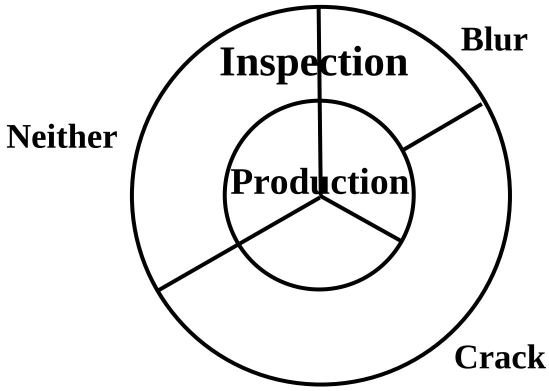


Figure 4-22: *Roulette-wheel representation of the frequencies and probabilities in Experiment 1 of Yamagishi (2003).*

The area between the inner and the outer circle depicts the proportion of removed and blurred gemstones due to inspection.

Since all cracked and half of the blurred gemstones have been removed, there remain 1/6 gemstones (out of all the gemstones produced) that are blurred and 1/3 gemstones that are neither blurred nor cracked.

Thus, the probability that within the set of inspected gemstones one is blurred is given by:

$$P\left(\text{blurred}\middle|\text{inspected}\right) = \frac{1/6}{1/6 + 1/3} = \underline{\underline{\frac{1}{3}}}$$

Tab. 4-17 reveals that the presentation of the roulette-wheel has a great impact on the proportion of correct answers. The probability format (frequencies vs. probabilities) has an effect, too. This is however the case only if no roulette wheel has been presented.

Tab. 4-17: *Percentage of correct responses as function of presentation of the roulette-wheel and of the frequency format.*

		Frequency	Probability
with	roulette-wheel	72.3%	69.8%
without	roulette wheel	42.5%	18.4%

Participants' improvement with the roulette-wheel diagram can be explained similarly as the improvement in case of frequency grids or outcome trees: The roulette-wheel enables the participant to conveniently access the relevant joint probabilities and joint frequencies, respectively.

The graphical method that was used to illustrate the relevance of base rates (cf. Ex. 4-52, on p. 226) can also be employed in the context of Bayesian reasoing.

Ex. 4-54: Monty Hall problem:

We assume that the candidate has chosen Door A, and that the host opens Door C (denoted by »C«). Figure 4-23 represents the hypotheses (concerning the location of the car) and the probabilities of the event »C« under the different hypotheses.

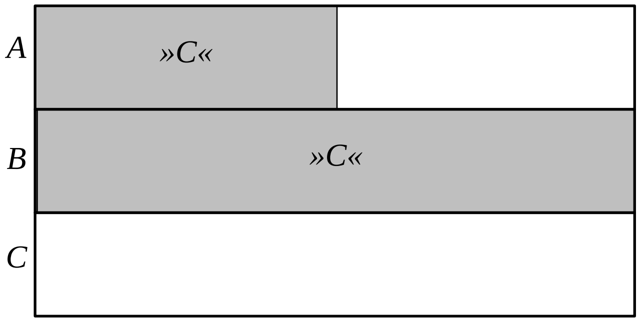


Figure 4-23: *A graphical representation of the Mont Hall problem.*

The size of the grey area if the car is behind Door A is only half the size of the grey area if the car is behand Door B. Thus the probability that the car is behind Door A, given the host opens door C is 1/3 (grey area under A divided by the total grey area).

From the problem solving perspective the following features render graphical representations and visual displays, respectively, favorable for various reasons:

1. A graphical representation provides a complete representation of the problem that includes the sets (or populations) involved, their embeddings, as well as their (relative) sizes.
2. In case of graphical aids used to improve Bayesian reasoning the joint probabilities $P(H_i \wedge E)$ of the hypotheses and the evidence are presented which simplifies the reasoning considerably since the operation of of combining the probabilistic information is not required for solving the problem.
3. The graphical representations also provides clear indications of how to compute the marginal probability $P(E)$ as well as the desired posterior probability $P(H_i|E)$ using different parts of the graph.

Recent studies reveal that the graphical methods are not necessarily superior to contingency tables representing joint probabilities if the graphical displays are presented only and not actively constructed by the participants (Böcherer-Linder, & Eichler, 2019; Talboy & Schneider, 2017). The issue of whether the active construction of the graphical representation provides an improvement over a more passive inspection of graphical displays is not well understood, however.

4.6.3 Formal Training

Various studies demonstrate a benifical effect of fromal statistical training on probabilistic reasoning (Fong, Krantz & Nisbett, 1986; Fong & Nisbett, 1991; Nisbett, Fong, Lehman, & Cheng, 1987). Two different types of evidence have been provided:

7. Training abstract statistical principles results in more and qualitative better statistical answers in different domains. In addition training statistical principles using examples from a specific domain transfers to problems from another domain. (Fong, Krantz, & Nisbett, 1986; Fong & Nisbett, 1991)
8. Higher education containing courses in statistical and methodological thinking increases the quality of methodological and statistical reasoning in students (Nisbett et al., 1987).

Concerning the first point, participants were trained using either examples from sports or from the topic of ability testing. For example, participants had to explain the fact that after the first two weeks of the major league baseball season, the top batter has typically an average of .450, yet no batter has ever had an average as that at the end of the season. Untrained persons usually provide exclusive causal explanations (that need not be necessarily wrong), whereas more sophisticated participants consider statistical answers, like small sample sizes and regression to the mean.It turned out, however, that the rate of transfer to a new domain decreases with delay of 2 weeks whereas the effect of training decreases only slightly for the trained domain. Nonetheless, even after a delay of 2 weeks trained participants performed better for the new domain than untrained control subjects.

With respect to the second type of evidence, indicated above, it turned out that for disciplines emphasing methodological and statistical training, like psychology and, to a lesser degree, medicine a substantial improvement of statistical and methodological reasoning (e.g. the importance of control groups or the problem of confounding) was observed after 2 years of education. In addition, students of psychology showed higher improvements than those of medicine. By contrast, students of chemistry and law showed no significant improvement in statistical and methodological reasoning after two years of education (At the beginning of their education, students from different disciplines exhibited about the same level of sophistication).

The evidence provided indicates that formal training in statistics and methodology results in improvments not only with respect to scientific reasoning but also for judgments concerning everyday problems.

4.6.4 Intelligence, Open-Mindedness, and Critical Thinking

Cognitive abilities, open-mindedness and the capability and motivate-on to critical thinking seem to be characteristics that are able to prevent people at least partially from committing biases and errors in judgment and reasoning tasks.
Stanovich & West (2008) found that people with higher mental abilities are less biased in their reasoning for some tasks but not for others. Tab. 4-18 provides an overview of biases that are correlated or not with cognitive ability.

Tab. 4-18: *Biases that are or are not affected by cognitive abilities (according to Stanovich & West, 2008, Table 8 on p.686).*

Biases not affected by cognitive ability	Biases affected by cognitive ability
Noncausal base rate use	Causal base rate use
Conjunction fallacy	Outcome bias
»Less-is-more« effect	Denominator neglect
Anchoring effects	Probability matching
Sunk-cost effect	Hindsight bias
Risk-benefit condounding	Ignoring $P\left(D\middle\|\bar{H}\right)$
Omission bias	Covariation detection
One-side bias	Belief bias in syllogistic reasoning
Certainty effect	Belief bias in modus ponens
WTP/WTA difference	Informal argument evaluation
Newcomb’s problem	Wason selection task
	Expectation maximization in gambles
	Overconfidence effect

The results in Tab. 4-18 reveal that greater cognitive abilities do not prevent people to commit fallacies, like base rate neglect in case of non-causal base rates (cf. Ex. 4-28, on p.150, and Ex. 4-29, on p.151) and the conjunction fallacy (cf. Ex. 4-11, on p.119). Stanovich and West (2008) explain this result by the fact that intelligence tests assess *cognitive capacity*, specifically working memory capacity. However, many fallacy are not due to a lack of cognitive capacity. Rather, they are caused by the inability to override plausible answers that have been been cued by the situation. These answers are based on representations missing important information that is relevant for a proper response due to a lack of knowledge and/or expertise: People either don’t know that the information is relevant or they are unable to apply their knowledge in the given situation. If the answer is satisfactory for the participant no further effort will be devoted to the problem and thus cognitive capacity is irrelevant.

This line of reasoning is confirmed by results of Toplak, West, and Stanovich (2011) who showed that performance for the Cognitive Reflection Test (CRT) (cf. Ex. 4-49, on p.216) is a better predictor of judgment and decision errors than traditional tests of intelligence. Not only does the CRT explain more variance than traditional intelligence items but it also exhibits *incremental validity*: If the CRT measure is included as a predictor in a regression analysis additionally to the measures of intelligence it explains additional variance that was not explained by the measures of intelligence. This indicates that judgments and decisions of more critical persons that are not satisfied by a plausible (but incorrect) answer coming to mind perform better.
Open-mindedness has been considered as a trait that might prevent people form committing judgment and decision biases (see, e.g., Baron, 2008). Open-mindedness refers to the capability to take on different positions and to perceive problems from different points of view. For example, in syllogistic reasoning tasks the number of alternative conclusions generated by participants was a better predictor of performance than intelligence (Newstead, Thompson, & Handley, 2002). Moreover, in case of myside bias (cf. Ex. 1-4 on p.8), open-mindedness has a positive effect in reducing the bias.

4.7 Criticisms of the »Heuristics and Biases« Program

In the eighties and nineties (of the twenty century) a number of criticisms of the heuristics and bias program arose. The criticism can be divided into three categories:

(1) A refusal to accept the normativity of axioms and definitions of the probability calculus.
(2) A criticism concerning the applicability of the probability calculus to reality.
(3) A criticism of methods and models used for modeling concrete problems.

The upshot of the criticism may be summarized as follows:

> *The observed errors and biases in probability judgments are due to the application of unjustified norms and methods. By consequence, the implications of the results concerning the assessment of human rationality are not justified.*

In the present section the tenability of these arguments is scrutinized.

4.7.1 Normativity of Axioms and Rules of the Probability Calculus

There are only few people that question the normativity of the axioms of probability theory. One of them was the British philosopher Jonathan Cohen (1923 - 2006). He claimed:

> *Nothing can count as an error of reasoning among our fellow adults unless even the author of the error would, under ideal conditions, agree that it is an error* (Cohen, 1981, p. 322).

Furthermore,

> *The intuitions of ordinary people are the basis for constructing a coherent system of rules and principles by which those same people can, if they so choose, reason much more extensively and accurately than they would otherwise do. Consequently these ordinary people cannot be regarded as intrinsically irrational in regard to any such cognitive activity* (Cohen, 1981, p. 322).

The two citations indicate that Cohen believes in a complete relativism with peoples' intuition being the ultimate authority for the assessment of the rationality of beliefs and actions. By consequence, Cohen assumes that the heuristics and biases approach has no implications for the assessment of human rationality. In addition, Cohen denies the normativity of Bayes rule even in cases were the probabilities involved are frequencies and, thus, the application of Bayes rule is completely unproblematic (cf. Krantz, 1981, for a criticism of Cohen's position).

Clearly, one is free to claim that everybody is rational unless he does not admit to have acted irrationally. However, it is highly questionable whether this position is of any value. If someone harms herself permanently because of her inconsistent system of beliefs then it seems useful to call this belief system irrational (clearly, everyone is free to use a different term like *sub-optimal* if he does not like the term *irrational*).

However Cohen (1981) is plainly wrong in claiming that the heuristics and bias approach has no implications with respect to the assessment of human rationality. There are at least two reasons for this: First, most people realize their erroneous intuition as soon as they learn more about the situation, and they agree to have performed an error. For example, Pinker (1997) reports that a student exclaimed spontaneously »I am ashamed of my species« after having learned about the conjunction error in the context of the Linda problem (cf. Ex. 4-11, on pp. 119).

Second, Stanovich und West (2000, 2008) have shown that for a number of biases people with increased cognitive abilities are less prone to judgmental errors and biases (cf. Tab. 4-18).

In general, most people accept the axioms and rules of the probability calculus as a sensible basis for consistent reasoning, similarly to logical rules being conceived of as relevant for consistent reasoning. However, the acceptance of the axioms and rules of probability does not imply that they are useful for the solution of complex problems of everyday life. This leads us two the second type of criticism: the issue concerning the applicability of the axioms and rules of probability.

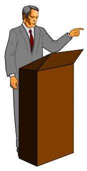
4.7.2 Applicability of the Axioms and Rules of the Probability Calculus

As detailed in the appendix the axiomatic conception of probability regards probability as a normed measure that is a generalization of measures like length, area, weight etc. This conception is clearly an idealization that does not apply exactly to real existing events.

A realization that comes quite close to the ideal are Casino games comprising long sequences where attention is paid that basic assumptions are generally met, like independence of events. By consequence, probability models describing the distributions of various events are a nearly perfect representation of reality. Note however that also in this case the models are only good approximations that do not capture reality perfectly. In general the following principle applies:

Principle 4-5: *Models of reality are generally incorrect*

Human models of reality do not represent their target domain perfectly. They are all more or less good approximations of reality and thus incorrect.



Comment 4-9: Applicability of measurement axioms:

The measurement axioms do not perfectly apply to the lengths of real objects. For example, it is not the case that the length of two pieces stuck together corresponds exactly to the sum of the single pieces since (due to the Pauli principle of quantum mechanics) it is impossible to stick two pieces together exactly.

In addition the order of the pieces stuck together makes a difference. This violates the axiom of commutative axiom according to which it is true that: $a + b = b + a$ (for arbitrary numbers a and b).

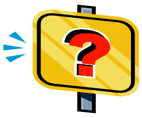
Due to the fact that models of reality are approximations only the issue of the applicability of a model amounts to the problem of whether the model approximates reality close enough in order to be useful.

Ex. 4-55: Adequacy of probability models (Continuation of Ex. 4-37, on p.175)

In Ex. 4-37 it was assumed that the probability of solving a problem by the student depends only on her general capacity of solve problems of this sort. In addition, it is assumed that this general capacity does not vary in course of solving the problems.

> The binomial model used to represent the probability of the number of problems solved is correct only it these assumptions are met. Clearly, in reality, these conditions are never fulfilled perfectly. However, for practical reasons it suffices if the assumptions are fulfilled approximately only.

The great utility of probability models has been revealed in many areas, like medical research, social research, insurance statistics, statistical mechanics, machine learning, etc. This indicates that many probability models are sufficiently good approximations to reality.
In the context of the heuristics and biases research two questions arise:



1. *Is it justified to use probability models for representing uncertain information in problems and scenarios used by Tversky and Kahneman?*
2. *Are the models used by Tversky and Kahneman adequate representations of the situations, or are they too simplified, neglecting important aspects of the situation?*

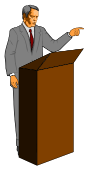
In the present chapter the first issue will be discussed in the context of the *conjunction fallacy*. The second question will be dealt with in the next chapter.
We first consider arguments that contradict an interpretation of the results of Tversky and Kahneman in terms of judgmental biases.

4.7.2.1 ARGUMENTS AGAINST THE INTERPRETATION OF THE RESULTS ON THE LINDA PROBLEM IN TERMS OF A CONJUNCTION FALLACY

Two arguments against an interpretation of the results for the Linda problem as a judgmental error have been forwarded:

1. *Ambiguity of the everyday conception of probability and violation of Grice's principles.*
   According to Hertwig and Gigerenzer (1999) lay persons interpret the term »probable« predominantly as »plausible« or »credible« since they do not know that they are confronted with a probabilistic problem task.
   Furthermore, the context suggests that the mathematical or axiomatic concept of probability (i.e. probability as a normed measure) cannot be relevant for the problem at hand since this would make the personal description of Linda irrelevant for solving the problem. The assumption that the axioms of probability theory make the presentation of the personal description of Linda irrelevant leads to the conclusion that the presentation of the personal description violates the conversational principle of relevance due to Grice, according to which irrelevant information should be avoided in a conversation. Thus people conforming to the Grice's maxims of conversation

have to reject the idea that axioms of probability theory are relevant for the Linda problem.



Comment 4-10:

The personal description of Linda is irrelevant since the critical relationship $P(B) \geq P(A \cap B)$ can be deduced from the Kolmogorov's axioms of the additivity of the probability measure for the union of disjoint events (as well as the non-negativity of probabilities):

According to basic set theory:

$$B = (A \cap B) \cup (\overline{A} \cap B),$$

where the events $A \cap B$ und $\overline{A} \cap B$ a disjoint (since A and $\overline{A}$ are disjoint). Consequently,

$$\begin{aligned} P(B) &= P[(A \cap B) \cup (\overline{A} \cap B)] \\ &= P(A \cap B) + P(\overline{A} \cap B) \\ &\geq P(A \cap B) \end{aligned}$$

The personal description is thus irrelevant for solving the problem.

Since people following the conversational maxims assume that the concept of probability is irrelevant they assume that different concepts, like plausibility or coherence, are relevant instead of the mathematical concept of probability.

Consequently, the »purported« conjunction fallacy is not really an error. It simply reflects the fact that people use different concepts than the investigators. In addition, according to Hertwig and Gigerenzer (1999) it might be more rational to follow the conversational norms and ignore the relationship of inclusion between the relevant sets.

2. *Applicability of the axioms of probability to subjective probabilities and probabilities of single events, respectively:*

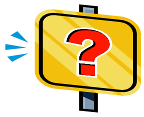
   According to Pinker (1997) the conjunction fallacy observed with the Linda problem (cf. Ex. 4-11, on p. 119) is not necessarily an error since the probabilities involved are subjective probabilities that may be interpreted as strengths of belief. Since mental entities are not extensional entities the application of relationships between sets (the latter are extensional entities) is not appropriate.

In conclusion both types of arguments boil down to the claim that the axioms of mathematical probability are not applicable in case of the Linda problem, and, by consequence, talking of a fallacy is inappropriate.

Let us now consider the soundness to these arguments.

4.7.2.2 Arguments Against the Counter-Arguments

The article of Hertwig and Gigerenzer (1999) raises two questions:



1. *Is their explanation correct that the cause of the conjunction fallacy is due to the fact that the situation violates the maxim of relevance?*
2. *Does this fact as well as the interpretation of »probable« as »plausible« turn the conjunction fallacy into a rational behavior?*

We now consider these two questions in detail:

1. *Violation of conversational maxims as the reason for the conjunction fallacy:*

There are two reasons that contradict this explanation:

i. Contrary to the claim of Hertwig and Gigerenzer the personal descriptions are relevant. They are required for the assessment of the probabilities of the other statements, for example, *Linda is a teacher in an elementary school*.

ii. The argument of Hertwig and Gigerenzer does not apply to the other cases of conjunction errors, like the example with Ronald Reagan (cf. Ex. 4-12 on p.121), since no personal description was presented in this case.

Comment:

Tversky & Kahneman (1983) used as an example of a conjunction error in the context of predictions a statement about the performance of Björn Borg who was the leading tennis player in 1980. Specifically the following two critical statements were contrasted:

B. Borg will lose the first set.
C. Borg will lose the first set but win the match.

72% of the participants committed the conjunction fallacy by estimating statement *C* as more probable than statements *B* which is lower than the rate in case of the Linda problem. However, the fact that nearly ¾ of the participants committed the fallacy despite the fact that no violation of Grice's maxims is present is in clear opposition to the argument of Hertwig and Gigerenzer.

For the Reagan example (cf. Ex. 4-12 on p.121) 68% of the participants committed the conjunction fallacy.

iii. This is further confirmed by a number of experiment of Sides, Oshershon and Bonini (2002) who tested the conjunction fallacy in a betting context where not mention was made of *probability*, *likelihood* or *chance*.

In this studies participants had to bet on one of two alternatives with one alternative containing the conjunction, e.g.

X: *The percentage of adolescent smokers in Texas will decrease at least 15% from current levels by September 1, 1999.*

$X \wedge Y$: *The cigarette tax in Texas will increase by $1.00 per pack and the percentage of adolescent smokers in Texas will decrease at least 15% from current levels by September 1, 1999.*

The two statements represented predictions at the time of testing and participants was offered are reward for correct betting. In addition care was taken, the *X* was not interpreted as $X \wedge \bar{Y}$.
In Experiment 1, 36 out of the 45 prticipants in the betting condition committed a conjunction error for at least one of the bets. The corresponding frequency for the standard condition involving probability judgments was 38 out of 45.

The conclusions following from the above considerations are clear-cut: Since neither the personal description is irrelevant for the Linda-Problem nor the fallacy vanishes in case of no violations of the maxims the explication of Hertwig and Gigerenzer (1999) is obviously defective.
Let us now consider the second question:

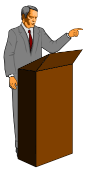
2. *Rationality of the conjunction fallacy:*

Hertwig und Gigerenzer (1999) conclude that the »presumed« conjunction fallacy is in fact a sign of social intelligence, and that it may be completely rational to commit the conjunction fallacy.
Now, there may be situations where it might be rational to commit a conjunction error, e.g. if it pays to make the opposite believe that one is quite stupid. However:

It is not rational, by any standard of rationality, to predict that a scenario with a lower chance of occurrence would have a higher likelihood of appearing compared to another one, if the latter is, in fact, at least as likely to occur as the first one.

This is the case for the Reagan and Bjorn Borg scenarios.

For this reason the claim of Hertwig and Gigerenzer concerning the rationality of the conjunction fallacy appears entirely unconvincing.



Comment 4-11:

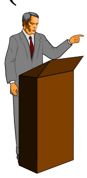
Please note that the argument that »probable« is interpreted as »plausible« goes in a similar direction as the argument of Cohen (1981) that the reasoning person is the ultimate instance for assessing the rationality, since whether a scenario is plausible or not depends on the subjective judgment.
If someone considers all the facts she might well judge the more probable scenario also as more plausible.

Let us now turn to the argument of Pinker (1997) according to which the axioms of probability are not applicable for the Linda problem since the probabilities involved refer to single events. The events are either true of false and, consequently, there does not exist any objective probability. Only subjective probabilities representing the strength of belief that an event will occur can be used.

This line of reasoning is questionable for the following reason: According to the classical view the probability of Linda being a bank teller and a bank teller that is active in the feminist movement, respectively, refers to the population of women that conform to the description of Linda (cf. the discussion in Section 4.3.1.6). With respect to this population there exists a certain probability that a member is a bank teller or that she is a bank teller and active in the feminist movement. If one further assumes that Linda is a randomly drawn subject from this population it is sensible to speak *of the probability that Linda has a specific characteristic* (e.g. that she is a bank teller).

A further objection to this argument may look like this: A person who assigns a higher probability to an event that is less likely to occur is, according to any standards, not consistent. By consequence, this person cannot maximize her expected subjective utility.

Finally, it should be mentioned that the subjective conception of probability is not the only one that can make sense of probabilities of single events [cf. the manuscript *Judgment and Decision Biases (Appendix Elements of probability theory)*]. Popper's (1959) propensity theory conceptualizes probabilities as characteristics of »probabilistic setups«, i.e. they can be seen as parameters that represent the whole configuration that produces the single events. This parameter can be estimated by using the outcomes of the setup. The precision of the estimation is clearly a function of the number of probabilistic experiments performed with the probabilistic setup. For example, if one plays Russian roulette the setup is determined by the number of bullets used, the revolver used, and how the cylinder of the revolver is spun. This all is part if the probabilistic setup. Obviously, it makes a great difference, also in the single case, whether there only one bullet (out of 6) in the cylinder than, say, 4.



Comment 4-12:

It seems that in the meantime Pinker has revised his opinion concerning the rationality of the conjunction fallacy. Now, he seems to regard it as a judgmental error (cf. Pinker 2003).

Let us summarize the previous discussion concerning the applicability of the axioms of probability:

1. The applicability of the axioms of probability requires that the domain of application meets a number of assumptions. In reality these assumptions are never met perfectly. However, the success of probability theory and statistics in different areas demonstrates that in these domains the axioms are met with a sufficient good approximation to render the application useful.

2. The claim that the conjunction fallacy is due to a conflict with the conversational principles of Grice is in no way convincing since the allegedly violations of the principle simply do not exist.
3. Likewise, the claim that the conjunction fallacy is an expression of rational behavior is untenable since persons committing the conjunction fallacy will, in general, provide more erroneous predictions.

Another line of criticism concerns the application of specific probability models as an adequate representation of the problem situation. This criticism will be addressed in the following chapter.

4.7.3 Normative Models Erroneously Applied

An argument concerning an erroneous application of Bayes theorem to the cab problem (cf. Ex. 4-33, on page 159) has been put forward by Birnbaum (1983), and has been reiterated by Gigerenzer and Murray (1987). This type of criticism may be summarized as follows:

> *The conditions under which the witness identified the color of the cab are different from the situation where the reliability of the witness was tested to correctly detect the colors of the cabs. Specifically, the witness adjusts his decision criterion according to the ratio of blue and green cabs. In the night when the witness identified the cab as blue the prior odds in favor of the blue cabs was 15/85. During the reliability test the odds was 50/50. As a result of the adjustment of decision criteria the hit rate and the rate of false alarms do not correspond to those given by Tversky and Kahneman.*

The main point of the criticism consists in the claim that the *values of the likelihoods* (representing the hit and false alarm rate) *vary with the base rates*. This is due to the fact that *the base rates exert an influence on the mechanism underlying the identification of the color of the cab.* It is well-known that base rates can in fact exert an influence on the identification of stimuli. The most famous model that can explain this fact and that was actually used by Birnbaum (1983) is the signal detection (SDT) model whose central features are shown in Figure 4-24. The SDT model is based on the following assumptions:

1. The decision of the witness to categorize a given color as blue instead of green is based on a continuous quantity that may be called »blueness«, i.e. the degree of perceiving the color as blue. This decision variable is a latent variable that cannot be observed directly. In Figure 4-24 this quantity is represented by the *x*-axis.
2. The presentation of stimulus (a color) does not result in exactly the same value on the internal scale »blueness«. Rather, the values of »blueness« are distributed with the location of the distribution depending on the color of the cab presented. In Figure 4-24 the distributions of the »blueness« are represented by the two normal density

curves: The green curve represents the distribution of »blueness« in case of a green cab being presented whereas the blue density curve represents the respective distribution if a blue cab is shown.
Please note that a green cab can result in a higher value of »blueness« than a green cab.

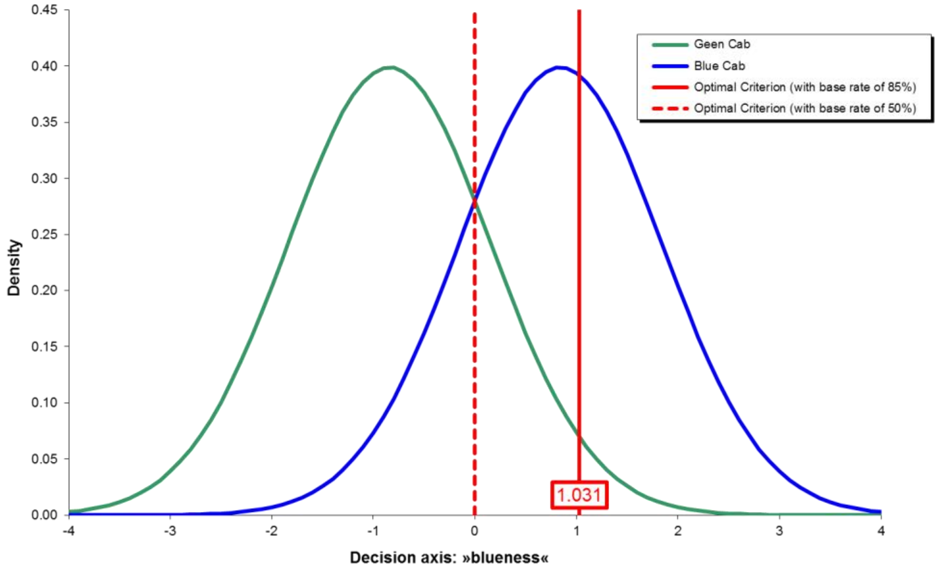


Figure 4-24: *Signal detection model for modeling the decision process of the witness for the cab problem: The green and blue curves represent the distribution of the decision variable in case of a green and blue cab, respectively, being presented. The dashed vertical line represents the optimal placement of the decision criterion in case of equal base rates of green and blue cabs. The full vertical line indicates the location of the optimal decision criterion in case of a ratio of 15/85 of blue vs. green cabs.*

3. To arrive at a decision the witness uses a decision criterion which separates the decision axis into two regions (The two vertical lines in Figure 4-24 indicate two possible decision criteria). If the »blueness« value of the cab presented is located to the left of the criterion a »green« response is provided, if it is located to the right the witness responds »blue«.
4. The witness has no control concerning his impression of the »blueness« of a cab. Thus he cannot change the form and location of the distributions of the values of the latent decision variable. On the other hand, he is able to adjust his decision criterion.

By adjusting his decision criterion the witness is able to maximize the probability of a correct response: In case of equal base rates of green and blue cabs the optimal criterion, i.e. the criterion that maximizes the probability of a correct response is located at the intersection of the two density curves. The dashed vertical line in Figure 4-24 indicates the respective location. In this case, the hit rate (of correctly identifying a blue cab as blue) is 80% which corresponds to the area under the blue curve to the right of the decision criterion. The false alarm rate (of erroneously identifying a green cab as blue) is 20%. This corresponds to the area under the green density curve to the right of the decision criterion. The percentage of a correct response is therefore 80%:

$$\begin{aligned} P(correct) &= P(»G« \wedge G) + P(»B« \wedge B) \\ &= P(»G«|G) \cdot P(G) + P(»B«|B) \cdot P(B). \\ &= 0.8 \cdot 0.5 + 0.8 \cdot 0.5 \end{aligned}$$

In case of knowing that the fraction of cabs is 85% the witness should shift his decision criterion to the right, in order to maximize the probability of a correct response. Consequently, he should emit more green (»G«) responses since there are much more green cabs. The optimal decision criterion for this case is shown by the full vertical line in. The optimal criterion results (for the distributions shown in Figure 4-24) in a false alarm rate of 3.1% and in a hit rate of 42.5% resulting in an overall rate of correct responses of 89%:

$$\begin{aligned} P(correct) &= P(»G«|G) \cdot P(G) + P(»B«|B) \cdot P(B) \\ &= 0.969 \cdot 0.85 + 0.425 \cdot 0.15 \end{aligned}.$$

On substituting the modified likelihoods, $P(»B«|B) = 0.425$ as well as $P(»B«|G) = 0.031$, into the Bayes formula:

$$\begin{aligned} P(B\,|»B«) &= \frac{P(»B«|B) \cdot P(B)}{P(»B«|B) \cdot P(B) + P(»B«|G) \cdot P(G)} \\ &= \frac{0.425 \cdot 0.15}{0.425 \cdot 0.15 + 0.031 \cdot 0.85} \end{aligned},$$

one gets a posterior probability of: $P(B\,|»B«) = 0.71$. This value is considerably closer to participants' modal estimate of 0.80 than the »correct« value of 0.41 that results from using Bayes theorem with the values provided by Tversky and Kahneman (cf. Ex. 4-33, on page 159).

This analysis demonstrates the following important aspect:

Modified base rates can result in a variation of the likelihoods in case of the base rates exerting an influence on the mechanism that determines the likelihoods.

From a modeling perspective the analysis presented results in an extension of the problem by *incorporating as an additional variable* that indicates whether the witness performed an adujstment of the criterion due to unbalanced base rates. Thus, the extended problem consists of three variables:

- *Color of the cab*: blue (B) vs. green (G),
- *Response of the witness:* »blue« (»B«) vs »green« (»G«), and
- *Adjustment of the decision criterion due to unequal base rates:* adjust (A) vs. do not adjust ($\bar{A}$).

Assuming stochastic independence between adjustment due to unequal base rates and the color of the cab, Figure 4-25 depicts the expanded cab problem.

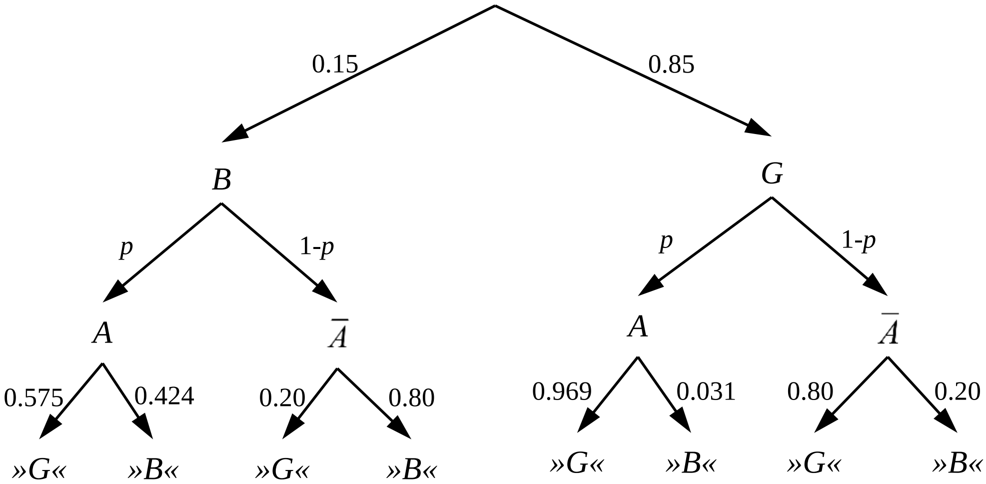


Figure 4-25: *Outcome tree of the expanded cab problem with the additional variable representing the adjustment of the decision criterion by the witness due to unbalanced base rates.*

The expanded representation of the problem elucidates that the relevant posterior probability is no longer the simple conditional probabilities $P(B|»B«)$ but the pair of probabilities $P\left(B|»B«,A\right)=0.71$ and $P\left(B|»B«,\bar{A}\right)=0.41$, depending on whether the participant adjusts for unequal base rates or not.

Note that the likelihoods of the expanded problem obey the following inequalities (cf. Figure 4-25):

$$P\left(»B«|B,A\right)\neq P\left(»B«|B,\bar{A}\right)$$
$$P\left(»B«|G,A\right)\neq P\left(»B«|G,\bar{A}\right)$$

This means that there is *no conditional independence* between the answer of the witness and the prior odds given the color of the cab since the prior odds may have an influence on whether the witness adjusts for unequal base rates or not.

The analysis presented raises two questions. Here is the first one:

Does this analysis do justice to the cab problem as formulated in Ex. 4-33, on page 159, i.e. is the analysis applicable to the cab problem?

The problem formulation states that: »*The court tested the reliability of the witness under the same circumstances that existed on the night of the accident* […] «. This indicates that the conditions were exactly replicated during the test which also comprises the different base rates of the cabs. By consequence, the analysis of Birnbaum (1983) is not applicable to the version of the cab problem of Ex. 4-33, on page 159.
In fact, the criticism of Birnbaum (1983) refers to an older version of the cab problem from the year 1980. In this version the test of the witness was described slightly different: The witness was tested under the same visual conditions as in the night of the accident. Whether the true base rates (85% green, 15% blue) had been used was not specified. Birnbaum (1983) assumed that during the test equally many blue and green colors were presented. Whether this is justified or cannot be decided. However, the argument of Birnbaum makes clear that for this older version the problem was not specified unambiguously.
To reiterate, the criticism is not justified for the version of Ex. 4-33, on page 159. This formulation of the problem indicates that the natural base rates were used during the test. This has been admitted by Birnbaum (1983) at the end of his article. However, this fact did not prevent Gigerenzer und Murray (1987) from replicating the analysis of Birnbaum without mentioning that the analysis does not apply to the most recent version of the problem.
It should be noted that Gigerenzer's (2001) criticism that content-blind norms are applied is also not justified since the description of the problem incorporates the known facts about the distribution of cabs as well as the performance of the witness to discriminate between colors.
The second question raised by the analysis is the following:

Do considerations about perceptual mechanisms as presented by Birnbaum influence the participants' thoughts during the process of solving the problem?

Or stated differently:

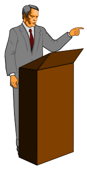
Are participants' responses influenced by complex considerations about the perceptual mechanism and the adjustment of decision criteria due to different base rates [Specifically, is the modal response of $P(B|»B«) = 0.80$ *due to such considerations]?*

It is quite reasonable that the participants regard the cab problem as a mathematical problem similar to other mathematical puzzles (like the Monty Hall problem). So, why should they doubt the specifications given in the problem formulation, and, instead, introduce highly questionable assumptions about perceptual mechanisms in order to come up with quite different likelihoods as those presented in the problem formulation [Note that the main reason of Birnbaum's analysis consists in questioning the probability information provided by the problem formulation].

A much simpler explanation of the erroneous probability judgments maintains that the participants simply do not understand the problem structure: People do not correctly represent the multiple ways to arrive at a result and that the probabilities of the outcomes due to these different ways have to be taken into account. Specifically, for the cab problem there are two different ways to induce a »blue» response of the witness: (i) The cab is blue and the witness identifies the color correctly [The probability of this outcome is $p = 0.15 \cdot 0.80 = 0.12$], and (ii) the cab is blue and the witness does not identify the color correctly [The probability of this outcome is $p = 0.85 \cdot 0.20 = 0.17$]. A real understanding of the problem has to incorporate a representation of these two possibilities to arrive at a »blue» response. Note also that the representation of the probabilities of the two ways reveals the importance of the base rates. The representation of the problems by means of outcome trees (cf. Ex. 4-35, on p.162) illustrates the two possible ways. As a result it provides a good means to explain the problem to lay persons (Sedlmeier & Gigerenzer, 2001).

If participants accept the probabilities given by the problem formulation then, clearly, the analysis of Birnbaum (1983) is irrelevant with respect to the issue of whether participants' performance is correct or not. An indication that participants take the values given by the problem formulation and do not make complicated and unjustified inferences is given by the results of an experiment by Gigerenzer und Hoffrage (1995). In their version of the cab problem, during the test the witness is positioned near the crossroads where the accident took place. This should guarantee that the same conditions are realized as at the time of the accident. As one might expect, this version resulted in no improvement indicating that participants are insensitive with respect to such details.

To summarize the discussion, the analysis of Birnbaum (1983) neither applies to the latest version of the cab problem nor is it of any significance with respect to the assessment of participants' judgmental errors.



Comment 4-13: Far-fetched arguments and unproved claims

In reviewing possible criticisms of heuristics and bias approach one gets the impression that the criticisms are, in part, based on far-fetched arguments. For example, the claim of Gigerenzer (2001) that one should not shift doors in the Monty Hall problem is far-fetched since it does not take account of the fact that the problem is a mathematical puzzle. The same is true for the arguments of Hertwig and Gigerenzer (1999) as well as Birnbaum (1983).

In addition the arguments of Hertwig and Gigerenzer (1999) and Birnbaum (1983) do not really apply. Consequently they provide no explanation of the judgmental fallacies.

Finally it should be noted that the authors do not take the effort to prove their claims, as is the case for Birnbaum (1983) or they are unable to provide a striking demonstration concerning the correctness of their argument, as in case of Hertwig and Gigerenzer (1999).

For example in the case of applying the signal detection model to the cab problem, Birnbaum (1983) as well as Gigerenzer and Murray (1987) assume that this model is a normative correct model that describes human performance correctly.

However, there are serious concerns whether decision criteria are really adjusted (cf. Balakrishnan 1999; Balakrishnan and MacDonald, 2002; Van Zandt, 2000). To my knowledge there exist no satisfying answers to these criticisms. Clearly, if the signal detection model provides no valid description of the witness' performance than the analysis of Birnbaum (1983) is useless.

Finally, it should be noted that a signal detection analysis was also performed in order to explain base rate neglect (cf. Ex. 4-28, on p.150) by Mueser, Cowan, and Mueser (1999). Similarly to Birnbaum (1983) they claim that the likelihoods resulting from the diagnostic information vary with the base rates.

However, similar to Birnbaum (1983) they are in no way interested to demonstrate empirically that the proposed mechanism is in any way relevant with respect to participants' behavior.

In addition, the criticism concerning the validity of signal detection models also applies to this analysis.

4.7.3.1 The Cab Problem and the Problem of the Proper Reference Classes

In Section 4.3.1.5 we discussed the problem of the non-monotonicity of probabilities and the problem of the proper reference class. The latter problem has also been discussed in the context of the cab problem by Levi (1981). He argues that the relevant reference classes of the cab problem are not the *green and blue cabs* but the *green and blue cabs involved in accidents*. It might well be possible that the blue taxis are committing more accidents and, by consequence, the ratio of 15/85 (of blue vs. green taxis) does not adequately describe the relevant prior odds that may be closer to 1.

There exists a fundamental principle that is invoked to handle ambiguous cases like the present one.

Principle 4-6: *Principle of insufficient reason / Principle of indifference:*

The principle of *insufficient reason* (*indifference*) states that equal probabilities must be assigned to each of competing assertions if there is no positive reason for assigning them different probabilities.

Source:

The Oxford Companion to Philosophy (2 ed.)

http://www.oxfordreference.com/view/10.1093/acref/9780199264797.001.0001/acref-9780199264797-e-2445#

With respect to the present case the principle of insufficient reason suggests that there is no reason to assume that the probability of committing an accident would be different for the green and blue taxis. Thus it is reasonable to assume equal probabilities for both types of cabs, and by consequence to accept the given prior odds.

However, Levi (1981) argues that the principle of insufficient reason can lead to inconsistencies. He, thus, rejects the solution provided by the principle and opts for ignoring prior probabilities altogether (which amounts to assuming prior odds equal to 1). The argument of Levi has a number of weaknesses:

First, to assume that the *green and blue cabs involved in accidents* proper reference classes is quite arbitrary. One might well assume that the proper reference classes are the *green and blue cabs involved in accidents in the night* or the *green and blue cabs involved in accidents in the night in a specific location in the town* (where the accident actually took place). In fact, using ones imagination, it is easy to generate further possible reference classes that may have some reasonableness.

Second, it is correct that the principle of insufficient reason can lead to inconsistencies (see, for example, Nickerson, 2004). However, is unclear how this might occur in the present case. Consequently, the general argument that the principle of indifference can lead to inconsistencies does not justify its non-application in the present case.

In fact, Levi (1981) does not provide any specific reason why the principle should not be applied in the present case.
The preceding discussion has shown that the arguments against various experiments of Tversky and Kahneman are not convincing. It should however be noted that this cannot be generalized to other cases. For example, with respect to the phenomenon called overconfidence there seems to exist a well-justified criticism of at least some of the studies (see, e.g. Juslin, 2001).

4.8 Exercises

Exercise 4-1: *Interpretation of probability*

Given: The following probability statement:

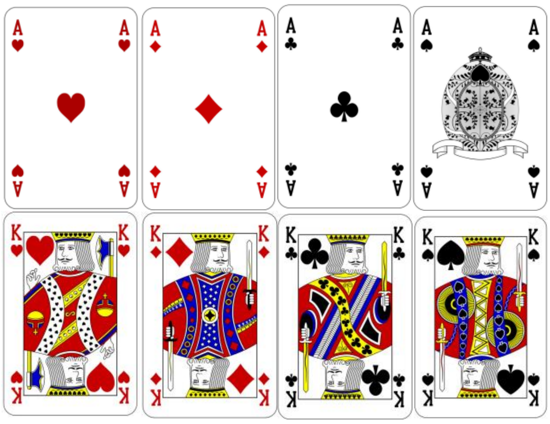
> *Consider a woman applying a home pregnancy test kit. If she tests positive, the probability that she is pregnant increases. Conversely, if the test is negative, the probability that she is pregnant decreases.* (Meder & Gigerenzer, 2014, p.128).

Please, give an interpretation of the statement applying:

(a) The objective conception of probability, and
(b) The subjective conception of probability.

Exercise 4-2: *The card problem of Copi (1968)*

Remove all cards except aces and kings from a deck, so that only eight cards remain, of which four are aces and four are kings.



From this abbreviated deck, deal two cards (at random).

(a) One of the two cards is an ace. What is the probability that the other card is an ace, too?

(b) One of the two cards is an ace of spades. What is the probability that the other card is an ace, too?

Hint: List all possible combinations of cards for two the different sub-populations involved in the two problems.

Exercise 4-3:

Given:

- ❑ Prevalence of a disease: 0.3%
- ❑ Sensitivity of a diagnosis: 90%
- ❑ Rate of false alarms: 3%

Let:
D = disease (present).
+ = positive diagnosis.

Explain, why $P(+|D)$ is higher than $P(D|+)$.

Hint: Show that the probability of a positive diagnosis is considerably higher than that of a disease.

Exercise 4-4:

Explain, why the formula of the conditional probability:

$$P(A|B) = \frac{P(A \wedge B)}{P(B)},$$

makes sense intuitively.
Hint: Which proportion is represented by the formula?

Exercise 4-5:

Please give an example of the asymmetry of conditional probabilities. Draw Venn diagrams and explain by means of these diagrams the reasons of one conditional probability being greater than the other one.

Exercise 4-6:

Consider a gamble with two (fair) dice that both are thrown once.

(a) What constitutes the population in this case?

(b) What is the probability of a double-six if the number of points shown by the two dice is greater than 10?

(c) What is the probability of at least one six if the number of point shown by the two dice is greater than 8?

Exercise 4-7:

The famous French mathematician Jean-Baptiste le Rond d'Alembert (1717-1783) believed that the probability of at least one head in two tosses of a fair coin is 2/3 (Nickerson, 2004).
He argued as follows: The sample space (the population) consists of three units:

- The event H (a head in the first throw, which ends the game since a head has turned up).
- TH (a tail (T) followed by a head), and
- TT (two tails)

Thus the probability of at least one head which conforms to the sum of the probabilities of the first two events is 2/3.

Please explain why this line of reasoning is flawed.

Exercise 4-8:

The dice problem of chevalier de Méré:
Which of the two events has a higher probability:

(a) To get at least one six in 4 tosses of a fair die?
(b) To get at least one double-six in 24 tosses of two dice?
(c) What are useful sample spaces (populations) for these two events?

Hint: Consider sequences of success and failures in each case.

Exercise 4-9:

Given:
Three hypotheses concerning the bias of a coin:

$$H_1 : \pi_1 = 1/4$$
$$H_2 : \pi_2 = 1/2$$
$$H_3 : \pi_3 = 3/4$$

Where π_i $(i = 1, 2, 3)$ denotes the probability of the coin landing *Heads*:

$$P(Heads|H_i) = \pi_i \text{ and } P(Tails|H_i) = 1 - \pi_i.$$

It is assumed that the three hypotheses represent the only possibilities, i.e. they cover the space of hypotheses.
In addition, assume that the prior probabilities for each of the three hypotheses are the same, i.e.:

$$P(H_1) = P(H_2) = P(H_3) = 1/3.$$

Now, the coin is tossed three time and the outcome E is *HTH* (*H* = *Heads*, *T* = *Tails*).
What is the probability of the three hypotheses given the observed outcome E: $P(H_i|HTH)$, (i = 1, 2, 3).

Comment: The single tosses are assumed to be independent.

Exercise 4-10:

Prove the following relationship:

$$P(A|B) > P(A|\bar{B}) \Leftrightarrow P(B|A) > P(B|\bar{A})$$

Hints:
1. Show, first, that

$$P(A|B) > P(A|\bar{B}) \Leftrightarrow P(A \wedge B) > P(A) \cdot P(B)$$

2. Keep the following relations in mind:

$$P(A) = P(A \wedge B) + P(A \wedge \bar{B})$$

$$P(B) = P(A \wedge B) + P(\bar{A} \wedge B)$$

$$P(\bar{A}) = 1 - P(A),\ P(\bar{B}) = 1 - P(B).$$

Exercise 4-11:

Prove the following relationship:

$$P(H|D) > P(H) \Leftrightarrow P(D|H) > P(D)$$

In words: The posterior probability of the hypothesis given the data is greater than the prior probability of the hypotheses if and only if the likelihood of the data given the hypothesis is greater than the probability of the data.

Hint: Inspect Bayes formula.

Exercise 4-12:

Given: The following exposition of the problem:

According to actual results it is know that a lie detector indicates (with perfect certainty) that the person is lying if the investigated person is lying.

The detector indicates that the person is lying in 50% of the cases when the person is not lying.

10% of the persons that are subjected to a lie detector test are lying whereas the rest is telling the truth.

Compute the probability that a person subjected to a lie detector test is lying if the detector indicates that the person is lying?

(i) Use an outcome tree to represent the problem and compute the required probability with the help of the outcome tree.

(ii) Compute the same probability using Bayes theorem in odds format.

Exercise 4-13:

Prove the following sequential property of Bayes theorem enabling the sequential updating of the probability of hypothesis H on the basis of two pieces of evidence, E_1 and E_2.

$$P(H|E_1, E_2) = \frac{P(E_2|H)}{P(E_2)} \cdot P(H|E_1)$$

under the following independence assumptions:

1. $P(E_2|H) = P(E_2|H, E_1)$, and

2. $P(E_2) = P(E_2 | E_1)$.

Thus, it is assumed that the second piece of evidence E_2 is conditionally independent from evidence E_1 (given hypothesis H) as well as unconditionally independent from evidence E_1.

Exercise 4-14:

Given:

A small company producing printing plates has three machines: a new Machine M_1, a slightly older one, M_2, and a very old one, M_3.

M_1 produces 5000 printing plates per day, M_2 produces 3000, and M_3 1000.

The error rates of the three machines are also different:

M_1 produces, on average, 1 erroneous plate per 300 pieces.

M_2 produces, on average, 1 erroneous plate per 100 pieces.

M_3 produces, on average, 1 erroneous plate per 70 pieces.

The executive producer selects randomly 1 printing plate from the set of plates produced on that day and recognizes that it was faulty. What is the probability that the plate was produced by M_1, M_2, and M_3, respectively?

(i) Use an outcome tree for representing the problem and for calculating the relevant conditional probabilities.

(ii) Compute the relevant probabilities using Bayes theorem in odds format.

Exercise 4-15: *Representation of the extended cab problem by means of an outcome tree and a contingency table based on probabilities:*

Given: The extended cab problem of Ex. 4-40 on page 188.

(i) Represent the problem by means of an outcome tree with proabilities.

(ii) Represent the problem by means of a contingency table with probabilities.

Exercise 4-16: *Bayesian reasoning problem in natural frequency format:*

Given:

The following Bayesian reasoning problem (Cascells, Schoenberger, & Grayboys, 1978):

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person's symptoms or signs.

State the problem in terms of natural frequencies.

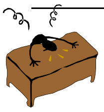
Exercise 4-17: *Testing for conditional independence:*

Use the data of to Tab. 4-19 to show that:

$P\left(»B_1«\middle|G\right) = P\left(»B_1«\middle|G, »B_2«\right) = P\left(»B_1«\middle|G, »G_2«\right)$.

Tab. 4-19: *Contingency table representing joint and marginal probabilities of the colors of the cab and the testimonies of the two witnesses for the extended cab problem.*

	Classification of the witnesses				
Color	»B1«, »B2«	»B1«, »G2«	»G1«, »B2«	»G1«, »G2«	Σ
B	72/1000	48/1000	18/1000	12/1000	150/1000
G	51/1000	119/1000	204/1000	476/1000	850/1000
Σ	123/1000	167/1000	222/1000	488/1000	1



Exercise 4-18: *Testing for the lack of conditional independence:*

Given:

The outcome tree of Figure 4-26 representing the extended cab problem. The symbols in the figure have the following meaning:

B = The cab is blue.
G = The cab is green.
$»B_1«$ = The first witness identified the cab as »blue«.
$»G_1«$ = The first witness identified the cab as »green«.
$»B_2«$ = The second witness identified the cab as »blue«.
$»G_2«$ = The second witness identified the cab as »green«.

Show that conditional independence of the identifications of the witnesses given the color of the cabs does not hold.

Hint:

Show, for example, that $P\left(»B_2«\middle|B, »B_1«\right) \neq P\left(»B_2«\middle|B\right)$.

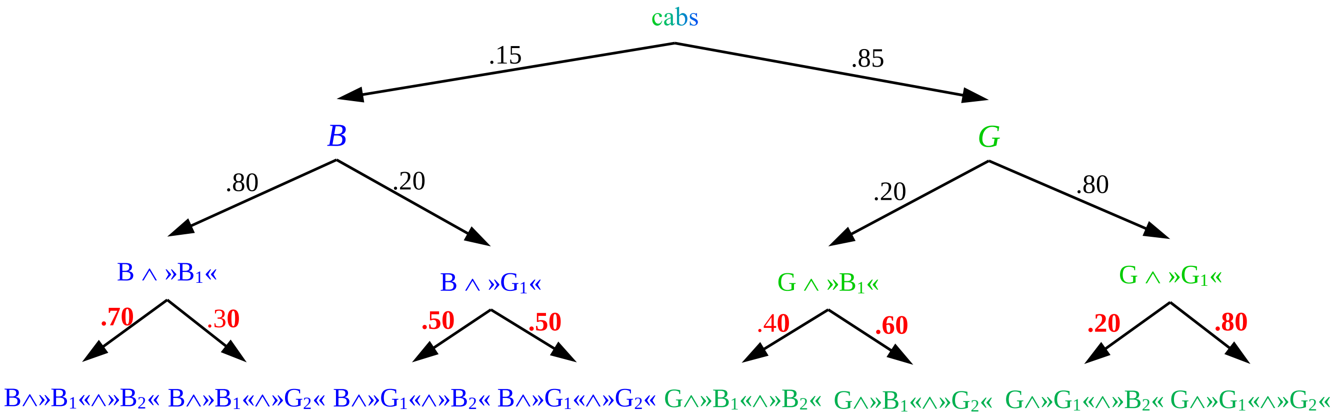


Figure 4-26: *Representation of the extended cab problem by means of an outcome tree in probability format. Conditional independence of the testimonies of the witnesses given the color of the cabs does not hold.*

Exercise 4-19: *Monty Hall Dilemma (Biased Monty):*

(i) Use an outcome tree to represent the Monty Hall problem of Ex. 4-47 (p. 207). Assume that the candidate has chosen Door *A* and that the host has no preference with respect to choosing Door *B* or *C*, i.e. he opens each door with the same probability of ½ in case of the car being located behind Door *A*.
Compute the probability that the candidate wins the car in case of the host has opened Door *C*, in case of no shift as well as in case of a shift.

(ii) Repeat the exercise under the assumption that the host has a preference for Door *C*, i.e. he always opens Door *C* if the car is not located behind this door.

Exercise 4-20: *Monty Hall Dilemma: Earthquake scenario*

Suppose you're on a game show, and you're given the choice of three doors (A, B, or C): Behind one door is a car; behind the others, goats. You pick a door, say Door A, and the host, who knows what's behind the doors, intends to open another door as an earthquake shatters the room and opens Door C, which has a goat.
Having recovered from the shock the host says to you, »Do you want to pick Door B?« Is it to your advantage to switch your choice?
Use an outcome tree to represent the problem and compute the relevant conditional probabilities. Is it to your advantage to switch doors?

Exercise 4-21: *Monty Hall Dilemma (Alternative conditioning event I)*

Assume that instead of the conditioning event that the host opens Door *C* we have the fact that the price is not behind Door *C* as the conditioning event.

Compute the conditional probability $P(A|\bar{C})$ of the price being behind Door *A* if the price is not behind Door *C*.

Exercise 4-22: *Monty Hall Dilemma (Alternative conditioning event II)*

Assume the setup of the Monty-Hall dilemma (Ex. 4-47, p. 207) with the following modification: Instead of the host opening Door *C* the candidate asks the host whether the price is behind Door *C*. The host answers truthfully either “yes” or “no”.

Compute the conditional probability $P(A|\mathit{Host\ answers\ »no«})$ of the price being behind Door *A* if the host answers »no« to the question of the candidate whether the price is behind Door *C*.

Exercise 4-23: *Monty Hall Dilemma with 4 Doors and 2 Cars*

Suppose you're on a game show, and you're given the choice of four doors (A, B, C or D): Behind two doors are cars; behind the other two, goats. You pick a door, say Door A, and the host, who knows what's behind the doors has to open two doors, one containing a goat and one containing a car. In addition, in case of a free choice the host selects doors randomly.

He opens Door B and D, behind B is a goat and behind D is a car, and asks you, »Do you want to pick Door C?«

Use a contingency table representation of the problem and compute the relevant conditional probabilities. Is it to your advantage to switch doors?

Hint: The events are combinations of doors.

Exercise 4-24: *The problem of the three cards*

Given:

Three cards:

1. A card with both sides being white.
2. A card with one side being white and the other one being black.
3. A card with both sides being black.

The cards are put into a box and mixed up. Then, one card is chosen at random and put on the table with one side up (The side is also chosen randomly). The side shown is white.

What is the probability that the other side is white (black)?
Hint: Draw an outcome tree.

Exercise 4-25: *The ball in the box*

Given:

A box contains a ball that is either white (W) or black (B). The probability of the ball being black is that same as the probability of the ball being white: $P(B) = P(W) = 1/2$.
A white ball is put into the box and the two balls are mixed up. Then a ball is chosen at random. It is white.
What is the probability that the other ball (in the box) is white (black)? Draw an outcome tree.

Exercise 4-26: *The tea-testing experiment*

A coworker of Sir Ronald Fisher claimed that she was able to recognize whether the milk has been poured into the tea or the tea has been poured into the milk. Fisher tested the assertion of the lady using a sample of cups containing tea with milk and the lady had to decide in each case whether the tea had been poured into the milk or the other way round (cf. Salsburg, 2001)
Assume that there is sample of 6 cups of tea with milk. Further, assume that we have three hypotheses:

H_1: The chance of the lady giving the correct answer is 1/2 (independently of whether the tea or the milk has been poured).

H_2: The chance of the lady giving the correct answer is 2/3 (independently of whether the tea or the milk has been poured).

H_3: The chance of the lady giving the correct answer is 5/6 (independently of whether the tea or the milk has been poured).

The prior probabilities of the three hypotheses are:

$$P(\mathrm{H}_1) = 1/2$$
$$P(\mathrm{H}_2) = 1/3$$
$$P(\mathrm{H}_3) = 1/6.$$

What is the probability of the three hypotheses if the lady got at least 5 correct answers out of the 6 attempts?

Hint: The likelihoods are binomial probabilities given by:

$$P(N \geq 5) = \sum_{n=5}^{6} \binom{6}{n} \cdot \pi^n \cdot (1-\pi)^{6-n},$$

where π represents the probability due to the given hypothesis.

Exercise 4-27: *The panda bears (McElreath, 2016, p.47)*

Suppose there are two species of panda bears: Both are equally common in the wild and live in the same places. They look exactly alike and eat the same food, and there is yet no genetic assay capable of telling them apart. They differ however in their family size. Species A gives birth to twins 10% of the time, otherwise birthing a single infant. Species B births twins 20% of the time, otherwise birthing singleton infants. Assume that these numbers are known with certainty, from many years of field research.

Now suppose you are managing a captive panda breeding program. You have a new female panda of unknown species, and she has just given birth to twins. What is the probability that her next birth will also be twins?

Hint:

The relevant events are:

(i) Species: A = Species A;
B = Species B.

(ii) Birth of twins in first birth:
T_1 = birth of twins in first birth;
$\overline{T}_1$ = birth of a singleton in first birth.

(iii) Birth of twins in second birth:
T_2 = birth of twins in second birth;
$\overline{T}_2$ = birth of a singleton in second birth.

The desired probability is: $P(T_2|T_1)$.

5. Biases and Paradoxes of Human Decisions

The present chapter presents empirical results as well as explanations concerning biases of human decision processes. Specifically, human decisions are affected by a number of influence factors that should be irrelevant for the decision. The present chapter is structured as follows: Section 5.1 presents empirical results concnering the most important decision biases, and effects on decision behavior, respectively. In Section 5.2 classical decision paradoxes are diecussed. Section 5.3 discusses the phenomenon of mental accounting that influences monetary decisions. Finally, in Section 5.4, Kahneman and Tversky's prospect theory and extensions of it are presented. This theory provides explanations of a number of decision biases.

5.1 Biases of Human Decision Making: Empirical Results

Research on decision processes revealed a number of factors that lead to non-optimal decisions. In the present section, some of the most important biases are discussed. In each of these cases, factors that are irrelevant for the decision in question exert an influence on the decision maker.

5.1.1 Framing effects

It is common wisdom that the way how different options are presented can exert a great influence on choices. Specifically, one might »frame« an outcome might be framed as either as a *gain* or as a *loss*. In the former case the outcome could be judged as more favorable as in the second case. As exhibited by the following example the tendency to frame outcomes in gains and losses can lead to somewhat strange results.

Ex. 5-1: Framing of outcomes (Jungermann, Pfister & Fischer, 2010, p. 235)

Here is an excerpt from the Margburger Magazine *Express* about the center for coronary heart diseases of children in Gießen (Germany):

> *Bis in die siebziger Jahre starben 20 Prozent der herzkranken Kinder in den ersten Lebenstagen. Fast jede grössere Familie hat so ein Kind auf dem Friedhof: Babys, die blau auf die Welt kamen, ein paar Tage nach Luft rangen und dann im Arm der Mutter starben. ‚Heute überleben achtzig Prozent‘ sagt Bauer mit leichtem Stolz.*

[Up to the seventies, 20 percent of the children with a heart disease died during the first days of their lives. One can find such a child on the cemetery for nearly every bigger family: babys who came blue into the world, struggling for breath for a few days and then dying in the arm of their mothers. 'Today eighty percent survive' Bauer explains with a certain element of pride]. [Translated by M.S.]

The classical study on the effect of framing on decision making was performed by Tversky and Kahneman (1981).

Ex. 5-2: Framing effects: The Asian disease problem (Tversky & Kahneman, 1981, p. 453)

Tversky and Kahneman provided the following two problems (more precisely, two versions of the same problem) to two groups of participants:

***Problem 1* [*N*=152]:**

Imagine that the U.S. is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people.

Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

If Program A is adopted, 200 people will be saved.

If Program B is adopted, there is 1/3 probability that 600 people will be saved and 2/3 probability that no people will be saved.

Which of the two programs would you favor?

72 percent of the participants decided in favor of Program A and 28 percent adopted Program B.

***Problem 2* [*N*=155]:**

The same cover story was used. However the alternatives were formulated differently:

Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimate of the consequences of the programs are as follows:

If Program C is adopted, 400 people will die.

If Program D is adopted, there is 1/3 probability that nobody will die, and 2/3 probability that 600 people will die.

Which of the two programs would you favor?

22 percent of the participants decided in favor of Program C and 78 percent adopted Program D.

The two versions the problems differ only with respect to whether the consequences are described in terms of gains (Problem 1) or in term of

losses (Problem 2). If the problem has been framed in term of gains, people were *risk aversive*, preferring the sure outcome, i.e., the saving of 200 people. If, on the other hand, consequences were framed as losses people were *risk seeking*.
Framing effects were also observed for medical experts.

Ex. 5-3: Framing effects and medical experts (McNeil, Pauker, Sox, & Tversky, 1982):

The 1153 participants consisted of three groups:

(i) 424 radiologists
(ii) 491 graduate students who had completed coursework in statistics and decision theory.
(iii) 238 ambulatory patients with different problems.

All participants received summary information on two forms of treatment for lung cancer: surgery vs. radiation therapy.
In approximately half of the cases, the summary information was framed in terms of cumulative probability of survival after a particular amount of time (e.g. 68% chance of living for more than one year).
In the other cases the summary information was cast in terms of mortality (e.g. a 32 percent chance of dying by the end of one year).
Since the danger of dying during or immediately after the surgery is the main disadvantage of this treatment it was hypothesized that surgery would be selected more often when summary information was framed in terms of probability of *living* than in terms of probability of *dying*.

Result: The hypothesis was confirmed:

- Surgery was preferred 75 percent of the time in the survival frame but only 58 percent in the mortality frame.
- The same pattern of results was found for each of the three groups of participants: radiologists, graduate students and patients.

Let us look at another important effect.

5.1.2 The Sunk Cost Effect

Concept 5-1: *Sunk Cost Effect (Fallacy):*

The *Sunk Cost Effect* consists in a tendency to continue an endeavor once an investment in money, effort, or time has been made. This tendency is based on the desire not to appear wasteful.

In case of the prior investment influencing the current decision despite the fact that it objectively should not influence it one can talk of a *fallacy*.

Let us look at a number of examples:

Ex. 5-4: Sunk Cost Effect (Arkes & Blumer, 1985):

Experiment 1:

Assume that you have spent $100 on a ticket for a weekend ski trip to Michigan. Several weeks later you buy a $50 ticket for a weekend ski trip to Wisconsin. You think you will enjoy the Wisconsin ski trip more than the Michigan ski trip. As you are putting your just-purchased Wisconsin ski trip ticket in your wallet, you notice that the Michigan ski trip and the Wisconsin ski trip are for the same weekend! It's too late to sell either ticket, and you cannot return either one. You must use one ticket and not the other. Which ski trip will you go on?

A. $100 ski trip to Michigan (N = 33)

B. $50 ski trip to Wisconsin (N = 28)

The number in brackets indicate the number of participants who chose the respective option. Thus, about 53% chose the less attractive Michigan trip. The prior investment seems to have influenced this decision.

This behavior is in clear opositoon to an axiom of traditional economic theory according to which decisions should be based on the costs and benefits that are expected to arise from the choice of each option.

Experiment 2 (Field Study):

Visitors of the Ohio University Theater who wanted to purchase a season ticket were offered a ticket. Season tickets were of three price classes: $15, $7, and $2.

Each visitor was randomly assigned to a specific price category. She did not know that there were different categories.

Result:

People with $15 tickets visited more plays in the first half of the season than those of the other two price categories.

Experiment 3:
Participants was presented one of the two scenarios:

Scenario A.
As the president of an airline company, you have invested 10 million dollars of the company's money into a research project. The purpose was to build a plane that would not be detected by conventional radar, in other words, a radar-blank plane. When the project is 90% completed, another firm begins marketing a plane that cannot be detected by radar. Also, it is apparent that their plane is much faster and far more economical than the plane your company is building. The question is: should you invest the last 10% of the research funds to finish your radar-blank plane?

Yes: $N = 41$

No: $N = 7$

Scenario B.
As president of an airline company, you have received a suggestion from one of your employees. The suggestion is to use the last 1 million dollars of your research funds to develop a plane that would not be detected by conventional radar, in other words, a radar-blank plane. However, another firm has just begun marketing a plane that cannot be detected by radar. Also, it is apparent that their plane is much faster and far more economical than the plane your company could build. The question is: should you invest the last million dollars of your research funds to build the radar-blank plane proposed by your employee?

Yes: $N = 10$

No: $N = 50$

The difference between the two scenarios consists in the fact that in Scenario A a lot of money had already been invested.
Note that there is no convincing economic reason to complete the project.

Experiment 4:
Same scenarios as in Experiment 3. People either got Scenario A or B and they had to judge the likelihood of a success for the respective scenario:

Scenario A: 41% success rate.
Scenario B: 34% success rate.

Experiment 5:
Same scenarios as in Experiment 3. However, for Scenario B instead of 1 Million Dollar, 10 Million Dollar could be spent on the project. This made no difference (in fact, exactly the same numbers as in Experiment 3 were obtained).

Cognitive Mechanism 5-1: Sunk cost effect and regret

The best explanation of the sunk cost effects is based on regret: The admission that one has wasted resources and the experience of the resulting sure loss produces regret.
To avoid these negative feelings, people are willing to spend more resources and to »throw good money after the bad«.

Experiment 6:

On your way home you buy a tv dinner on sale for $3 at the local grocery store. A few hours later you decide it is time for dinner, so you get ready to put the tv dinner in the oven. Then you get an idea. You call up your friend to ask if he would like to come over for a quick tv dinner and then watch a good movie on tv. Your friend says "Sure."
So you go out to buy a second tv dinner. However, all the on-sale tv dinners are gone. You therefore have to spend $5 (the regular price) for the tv dinner identical to the one you just bought for $3. You go home and put both dinners in the oven.
When the two dinners are fully cooked, you get a phone call. Your friend is ill and cannot come. You are not hungry enough to eat both dinners. You can not freeze one. You must eat one and discard the other. Which one do you eat?

$3:	$N = 2$
$5:	$N = 21$
No preference:	$N = 66$

Since both dinners are identical there is no reason to prefer one dinner over the other one.
With respect to regret things are different since to waste $3 produces less regret than wasting $5.

Experiment 7:

This experiment demonstrates that a future investment can be influenced by whether thoughts about waste are elicited or not:

Scenario A.
As the owner of a printing company, you must choose whether to modernize your operation by spending $200,000 on a new printing press or on a fleet of new delivery trucks.

You choose to buy the trucks, which can deliver your products twice as fast as your old trucks at about the same cost as the old trucks.

One week after your purchase of the new trucks, one of your competitors goes bankrupt. To get some cash in a hurry, he offers to sell you his computerized printing press for $10,000. This press works 50% faster than your old press at about one-half the cost.

You know you will not be able to sell your old press to raise this money, since it was built specifically for your needs and cannot be modified. However, you do have $10,000 in savings. The question is should you buy the computerized press from your bankrupt competitor?

Yes: $N = 49$
No: $N = 15$

Scenario B.

As the owner of a printing company, you must choose whether to modernize your operation by spending $200,000 on a new printing press or on a fleet of new delivery trucks.

You choose to buy the press, which works twice as fast as your old press at about the same cost as the old press.

One week after your purchase of the new trucks, one of your competitors goes bankrupt. To get some cash in a hurry, he offers to sell you his computerized printing press for $10,000. This press works 50% faster than your old press at about one-half the cost.

You know you will not be able to sell your old press to raise this money, since it was built specifically for your needs and cannot be modified. However, you do have $10,000 in savings. The question is should you buy the computerized press from your bankrupt competitor?

Yes: $N = 43$
No: $N = 38$

Note that buying the new press would result in the same proportion of improvement in both scenarios. However people are less willing to make the investment in Scenario B. The typical argument went like this: “I already have a good, new press that costs a lot of money.”

Thus, the new investment renders the old one as wasteful resulting in a greater tendency to reject it.

Experiment 8:

This experiment demonstrates that personal involvement increases the sunk cost effect.

The experiment uses the same scenario as Exp. 3, however with an important modification: It is not the participant as president of the company who has to make the decision. Instead the company is described as a third person and the participant should advice the company whether it should do the investment of the residual money.

This resulted in a significant reduction of positive answers for the sunk cost Scenario A:

Yes: *N* = 37 (Exp. 3: *N* = 41)
No: *N* = 21 (Exp. 3: *N* = 7)

A possible explanation of this effect of personal involvement may be that the experienced regret is higher in case of personal involvement.

Experiment 9:

This experiment investigates whether the sunk cost effect vanishes if no previous investment has been made.

The scenario of Experiment 1 was used with the modification that both trips (to Michigan and to Wisconsin) have been won in a lottery. The participant who prefers to go to Wisconsin later finds out that the ski trip to Michigan is worth $50 and that to Michigan $100. The following numbers were obtained:

$100 ski trip to Michigan: *N* = 44 (Exp. 1: *N* = 33)
$50 ski trip to Wisconsin: *N* = 42 (Exp. 1: *N* = 28)

There is no significant difference between Exp.1 and Exp.9. Obviously, even if no direct costs are involved the effect of wasting is present.

Experiment 10:

This experiment demonstrates that more sophisticated students with respect to economics are not less prone to the sunk cost effect. Experiment 1 ($100 ski trip to Michigan vs. $50 ski trip to the preferred Wisconsin) was administered to the students. Tab. 5-1 exhibits the result for the three groups of students:

Tab. 5-1: *Number of Selections of Options by Different Groups of Students in Experiment 10.*

	Economics students	Psychology students with no economics	Psychology students with economics
$100 trip	20	22	19
$50 trip	41	39	40

Comments:
No significant difference was found between groups.
For these students the fraction of participants selecting the $100 trip is significantly less than that found in Exp. 1.

The sunk cost effect may be influential for different types of decisions (not only financial ones), for example:

- Should I continue an unhappy relationship in which I have already invested so much?
- Should I continue this unhappy job (University study)? I have already spent so much time and effort.
- Should I leave the movie which is not enjoyable at all? But I have spent Sfr. 15.- as entrance fee.
- Should I make the trip for which I have already paid despite the fact that I feel sick?
- Ending the Vietnam War: It was argued that this would be a waste of lives.

The sunk cost effect may also be used in political arguments, for example to help lancing big projects with high investments.

Ex. 5-5: Construction of a nuclear power plant

In the seventies of the twenties century the Austrian government decided to build an atomic power plant. Since there existed a considerable resistance against this project in the population and the election period was approaching, the government decided to allow for public vote.

However, they decided that the vote should take place only after the construction work of the power plant had been finished.

From an economic perspective, the decision to enable the vote only after having finished the construction is difficult to justify since in case of a rejection additional millions will be wasted.

Taking the sunk cost effect into consideration, the decision to delay the vote after the end of erection makes sense due to the increased waste of resources and the resulting regret. This might increase the chances that people would vote in favor of the project.

Comment:
The policy did not work however, and the project was rejected.

5.1.3 Independence from Irrelevant Alternatives

It seems intuitively plausible that adding an unattractive alternative to an existing set of alternatives should have no great influence on the relative choice probabilities with respect to the other (more attractive) options. However, peoples' choice behavior exhibits this sort of influence. Here are two examples:

Ex. 5-6: Attraction effect (Simenson & Tversky, 1992)

One group of subjects had the choice between receiving $6 or a nice Cross pen. This resulted in the following percentages of choices:

$6:	64% chose this option.
Cross pen:	36% chose this option.

A second group received also the two options and, in addition, a third one consisting of a less nice Cross pen.

$6:	58% chose this option.
Cross pen:	46% chose this option.
Less attractive Cross pen:	2% chose this option.

Obviously, the addition of a less attractive pen increased the attractiveness of the other pen (instead of decreasing it). Note that the less attractive pen was practically never chosen.

Here is another example of a violation of the principle of independence from irrelevant alternatives

Ex. 5-7: Compromise effect (Simenson & Tversky, 1992)

Given: 3 possible options

A: A camera of high quality and price.
B: A camera of intermediate quality and price.
C: A camera of low quality and price.

If the choice set consisted of Options *A* and *B* people chose both options about equally often. Adding Option *C* the set resulted in higher choice rates for Option *B* compared to Option *A*.

Thus, despite the fact that Option *C* is itself quite unattractive it changes the decision behavior for the other options since Option *B* is now considered as a compromise between the exensive high-quality camera *A* and the cheap low-quality camera *C*.

Both example illustrate that a completely unattractive alternative may well influence the choice behavior with respect to the other, more attractive, alternatives.

5.1.4 The Endowment Effect

Concept 5-2: *Endowment Effect:*

The *Endowment Effect* consists in peoples' tendency to value object higher if they own it. As a result, people's maximum willingness to pay (WTP) to acquire an object is usually lower than the least amount they are willing to accept (WTA).

5.1.5 Relative versus Absolute Savings

The utility of saving of a certain amount of money should be independent of the total amount of money to be spent in a given situatain. This means that a saving of, say, Sfr. 100 should have the same value independently of whether the total amount of money spent is Sfr. 1,000 or Sfr. 10,000.

Ex. 5-8: Relative versus absolute savings (Thaler, 1980):

Participants received one of the following two versions of a decision problem.

Version 1:

Imagine that you go to purchase a calculator for $30. The calculator salesperson informs you that the calculator you wish to buy is on sale for $20 at the other branch of the store which is ten minutes away by car.
Would you drive to the other store? Option A: Yes, Option B: No.

Version 2:

Imagine that you go to purchase a jacket for $250. The jacket salesperson informs you that the jacket you wish to buy is on sale for $240 at the other branch of the store which is ten minutes away by car.
Would you drive to the other store? Option A: Yes, Option B: No.

Result: Significantly more participants were willing to drive to the other store in Version 1 than in Version 2.

The results can be explained by means of prospect theory's utility curves for losses (cf. Figure 5-1, p. 275): Initial losses with respect to the status quo are experienced as more painful than additional losses where a certain amount of loss has already occurred. Thus to loose Sfr. 1,000 is experienced as more unpleasant compared to a loss of additional Sfr. 1,000 if Sfr. 10,000 had already been lost.

Peoples' tendency to focus on relative savings is capitalized by salespersons in case of greater investments. For example, carsellers tend to offer additional facilities that are relatively cheap compared to the basic price of the car.

5.1.6 Less is More Effect

Ex. 5-9: Less is more effect in different groups of cognitive ability (Stanovich & West, 2008):

Participants receive one of three possible version of the following problem:

Form A:

Imagine that highway safety experts have determined that a substantial number of people are at risk of dying in type of automobile fire. A requirement that every car have a built-in fire extinguisher (estimated cost, $300) would save the 150 people who would otherwise die every year in this type of automobile fire. Rate the following statement for yourself: I would be extremely supportive of this requirement.

Form B:

The term *150 people* was replaces by *98% of 150 people.*

Form C:

The term *150 people* was replaces by *98% of 150 people.*

Participants answered on a 6-point scale:
1 = strongly disagree
6 = strongly agree

Results:

Tab. 5-2 indicates that the support in favor of the policy is higher for the less effective versions of the policy. In addition, this tendency is slightly (but not significant) stronger for participants with high cognitive ability.

Tab. 5-2: *Degree of support in favor of a policy as a function of its effectiveness and the cognitive ability of participants.*

	Form		
SAT	A: 150	B: 98% of 150	C: 95% of 150
Low	4.44	4.62	4.61
High	4.20	4.82	4.92

5.2 Paradoxes of Human Decision Making

5.2.1 The Allais-Paradox

Ex. 5-10: Certainty effect and Allais paradox (Kahneman (2011, page 384):

Consider the following two decision problems. How would you decide?

Problem I:

Option A	A gain of Sfr. 520 000.- with probability p = 0.61 or winning nothing with probability p = 0.39.
Option B	A gain of Sfr. 500 000.- with probability p = 0.63 or winning nothing with probability p = 0.37.

Problem II:

Option A'	A gain of Sfr. 520 000.- with probability p = 0.98 or winning nothing with probability p = 0.02.
Option B'	A gain of Sfr. 500 000.- with probability p = 1.00.

Result:

Most people choose Option A in *Problem I* and Option B' in *Problem II*.

Analysis:

These preferences are inconsistent. If one prefers Option A in *Problem I* then she should prefer Option A' for *Problem II*.

One can illustrate the chances of winning in *Problem I* by means of two urns, A and B, representing the two options:

- Urn A contains 61 balls labeled 520,000 and 39 balls labeled 0.
- Urn B contains 63 balls labeled 500,000 und 37 balls with the label 0.

A ball is drawn randomly from the Urn selected by the decision maker.

Now, the two options of *Problem II* result from those of *Problem I* by replacing 37 balls of Urn A labeled 0 by 37 balls labeled 520,000 (resulting in Urn A’), and by replacing 37 balls of Urn B labeled 0 by 37 balls labeled 500,000 (resulting in Urn B’).

Obviously, the modification of Urn A is more valuable than that of B since for both the same number of balls has been replaced but the values of the single balls are higher for A than B: 520‘000 vs. 500’000.
Consequently, preferring A in *Problem I* should be accompanied by preferring A’ in *Problem II*.
Interpretation:
The important reason for the observed inconsistency of peoples choices consists in the fact that Option B‘ in *Problem II* permits a sure (high) gain. For this gain people are willing to accept a lower return. This is called a *certainty effect*.
Peoples’ decisions are driven by their desire by to avoid *regret* (and anger) that would result in case of a negative outcome for Urn A’.

Allais paradox contradicts the Subjective Expected Utility (SEU) theory (cf. Concept 1-1, on p. 3), which is easily demonstrated. Peoples’ preferences for *Problem I* can be represented by the inequality:

$$0.61 \cdot u(\text{Sfr. } 520{,}000) > 0.63 \cdot u(\text{Sfr. } 500{,}000). \qquad (5\text{-}1)$$

Where $u()$ represents the subjective utility of the respective amount of money. The equation can be transformed to:

$$\frac{u(\text{Sfr. } 520{,}000)}{u(\text{Sfr. } 500{,}000)} > \frac{0.63}{0.61} \quad (=1.03). \qquad (5\text{-}2)$$

On the other hand, peoples’ preferences for *Problem II* can be represented by the inequality:

$$0.98 \cdot u(\text{Sfr. } 520{,}000) < 1.00 \cdot u(\text{Sfr. } 500{,}000) \qquad (5\text{-}3)$$

or:

$$\frac{u(\text{Sfr. } 520{,}000)}{u(\text{Sfr. } 500{,}000)} < \frac{1.00}{0.98} \quad (=1.02). \qquad (5\text{-}4)$$

Obviously, both inequalities cannot be true at the same time.

Ex. 5-11: Sicherheitseffekt und Vorsorge gegen Gefährdungen:

Der Sicherheitseffekt spielt auch im Kontext des Kaufes bzw. Verkaufes von Versicherungen eine Rolle. Personen präferieren Versicherungen, welche sie perfekt schützen.

Dies hat einerseits zur Folge, dass viel zu hohe Prämien bezahlt werden, nur um einen perfekten Schutz zu erhalten.

Beispielsweise zahlen Schweizer zu hohe Krankenversicherungsgebühren, weil sie einen viel zu geringen Selbstbehalt nehmen.

Andererseits wird oft übersehen, dass Versicherungen nur einen Teil der möglichen Schäden abdecken: Eine Feuerversicherung schützt nur bei Brandschäden, eine Sturmversicherung nur bei Schäden durch Unwetter etc.

Ein schöne Demonstration des Sicherheitseffekts im Kontext des Schutzes gegen Gefährdungen stammt von ….

5.3 Mental Accounting

Let us start with three examples from Thaler (1985) demonstrating our consumer behavior.

Ex. 5-12: Consumer behavior (Thaler, 1985)

Example 1:
Two couples went on a fishing trip in the northwest and caught some salmons. They packed the fish and sent it to home on an airline.
However, the fish was lost in transit and they received $300 from the airline. The couples take the money, go out to dinner, and spend $225. They never spent that much in a restaurant before.

Example 2:
Mr. *X* admires a $125 cashmere sweater at the department store. He declines to buy it, feeling that it is too extravagant. Later that month he receives the same sweater from his wife for a birthday present. He is very happy.
Mr. and Mrs. *X* have only joint bank accounts.

Example 3:
Mr. and Mrs. *Y*. have saved $15000 toward their dream vacation home. They hope to buy the home in five years. The (saved) money earns 10% in a money market account.
They just bought a new car for $11000 which they financed with a three year-car loan at 15%.

The three examples illustrate that people behave as if money has little labels attached to it indicating to where it belongs:

- ❑ In Example 1 the $300 are put in the categories »unexpected income« and »food«. As Thaler (1985) notes, the extravagant dinner would not have occurred if each couple had received a yearly salary increase of $150.

- ❑ In Example 2 the spending is registered under the label »gift«. People tend to give as gifts items that they would not buy themselves.
- ❑ In Example 3 the couple wants to keep apart the money labeled »dream vacation home« from the rest of their money with obvious economic costs.

5.4 Prospect Theory

Kahneman & Tversky (1979).

5.4.1 Utility Functions

Utilty is evaluated with respet to a reference point representing the current status-quo.

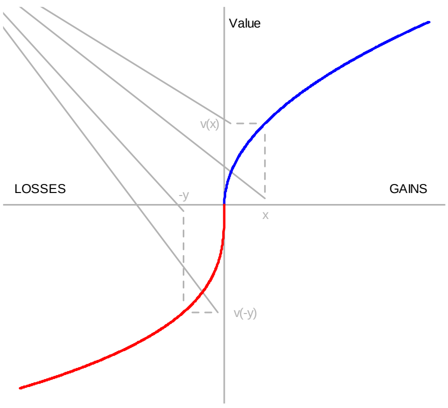


Figure 5-1: *Form of the value functions proposed by prospect theory: The value function for gains is concave (curved upward) whereas the value function for losses is convex (holds water). The value function is steeper for losses.*

5.4.2 Subjective Weighting of Probability Information

Most
Die meisten Personen überschätzen kleine und überschätzen hohe Wahrscheinlichkeiten.
Certainty effect. Pseudocertainty (Slovic, Fischoff, & Liechtenstein (1982) Plous p. 100)

Die ist in Abb. 5-1 illustriert, welche die subjektive Wahrscheinlichkeiten in Abhängigkeit von der objektiven darstellt.

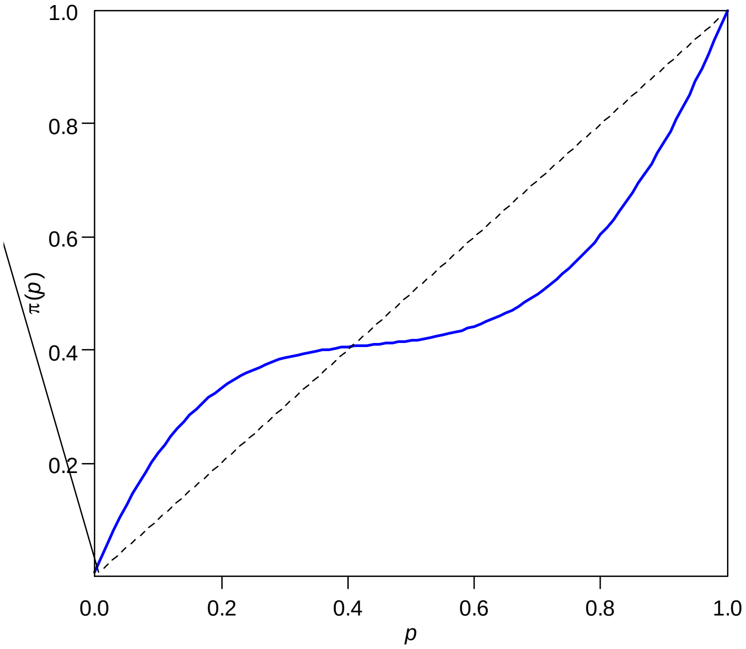


Abb. 5-1: *Subjektives Gewicht $\pi(p)$ als Funktion der aktuellen Wahrscheinlichkeit p (nach Tversky & Kahneman, 1992).*

Die Überschätzung von kleinen Wahrscheinlichkeiten mag mit verantwortlich dafür sein, dass Leute an Lotterien teilnehmen mit astronomisch geringen Gewinnwahrscheinlichkeiten (von mehr als 10^{-8} für den Hauptgewinn).

Eine weitere Folge besteht in der Tendenz, ungünstige Versicherungen zu akzeptieren. So haben z.B. die meisten Finanzprodukte mit Versicherung eine schlechte Performance (abgesehen davon, dass – wie im Zuge der Lehman-Pleite offensichtlich wurde – Versicherungen auch keinen perfekten Schutz bieten, was jedoch die meisten Anleger nicht wussten bzw. ignorierten).

Die Tendenz der Überschätzung kleiner Risiken ist vor allem für Personen in leitenden Positionen (Politiker, CEOs) bedeutsam: *Die Reduktion eines Risikos um – sagen wir 1% – hat grundsätzlich den gleichen Nutzen, unabhängig davon, ob es sich um eine Reduktion von 50% auf 49% handelt oder um eine Reduktion von 1% auf 0%.* Oftmals ist jedoch erstere Reduktion mit geringeren Kosten verbunden als letztere. Daher kann man oft mit viel geringerem Aufwand eine Verbesserung erzielen, indem ein mässig hohes Risiko weiter verringert wird anstatt ein kleines Risiko auf Null zu reduzieren.

5.5 Probability Matching (PM) and Decision Diversification

I, first, explain the phenomenon of probability matching (PM) and why the strategy of probability matching is not optimal. This is followed by a presentation empirical results concerning PM and various explanations of the phenomenon and the closely related phenomenon of decision diversification.

5.5.1.1 The Phenomenon of Probability Matching

Cognitive Mechanism 5-2: Probability Matching (PM)

Probability matching concerns the tendency of humans and animals to mirror the probability distributions of presented stimuli in their response probabilities.
In general, the strategy of probability matching results in non-optimal behavior.

The phenomenon of probability matching was investigated intensively in the context of experiments on learning in the sixties (of the twentieth century). The setup of a typical experiment is as follows: There are two types of stimuli: A red lamp that turns on with a probability of .70, and a green lamp that flashes up with a probability of .30. The task of the participants consists in predicting which light will turn on. Each response is followed by a feedback consisting in turning on the respective light.
For most participants the proportions of predictions reflect the actual probabilities of the two lights being turned on, i.e. in 70 percent of the trials they predict that the red light will turn on and in 30 percent of the trials they predict that the green light will flash up.
The strategy of probability matching is not the optimal strategy with respect to the minimization of the prediction error. The optimal strategy consists in the (constant) choice of the event with the highest probability of occurrence. In the actual example participants should always answer that the red light will be turned on.
The fact that probability matching is a suboptimal strategy can be easily demonstrated by reference to the actual example:
Tab. 5-3 shows the (expected) joint probabilities of the events: (a) *experimental outcome* und (b) *participant's response*, with the probabilities used (70% [red light] vs. 30% [greed light]) in case of probability matching being applied.

	Participant's prediction		
Outcome	»Red light«	»Green light«	Σ
Red light	0.49	0.21	0.70
Green light	0.21	0.09	0.30
Σ	0.70	0.30	

Tab. 5-3: *Joint probabilities of the two variables experimental outcome and participant's predictions.*

The computation of the probabilities in the Tab. 5-3 is based on the following consideration: The experimental outcome and participant's prediction are stochastically independent since the former is a pure random process. By consequence, the joint distribution results from the product of the marginal probabilities of the two events [Cf. Appendix, Section 2.1.3.1, Equation (2-3)]:

$$P\left(Outcome, Prediction\right) = P\left(Outcome\right) \times P\left(Prediction\right)$$

For example,

$$P\left(Outcome = red\ light,\ Prediction = "red\ light"\right) = 0.7 \times 0.7 = 0.49\,.$$

A correct response consists in an agreement of experimental outcomes with predictions. This occurs in the following two situations: Either the red light turns on and the participant had predicted the red light or the green light turns on and the participant had predicted the green light. Since both events are disjoint the probabilities can be added, and on gets:

$$\begin{aligned} P\left(Response\ correct\right) &= P\left(Outcome = red, Prediction = "red"\right) \\ &\quad + P\left(Outcome = green, Prediction = "green"\right) \\ &= 0.49 + 0.09 = 0.58 \end{aligned}$$

The probability of a correct response using PM (p = .58) is thus distinctly lower than the probability using an optimal strategy of always choosing the event with the highest probability: p = .70. Thus, PM is a sub-optimal strategy. Let us now consider some empirical phenomena concerning PM.

5.5.1.2 Empirical Findings concerning Probability Matching

There are a number of different findings throwing some light on the nature of PM.

5.5.1.2.1 PM and decision diversification

A phenomenon closely relate to PM is called decision diversification. This consists in the fact that with repeated choices people do not always choose the same (optimal) alternative but tend to choose different alternatives (they diversify their choices). A simple experiment of Rubinstein (2002) illustrates the phenomenon.

Ex. 5-13: Decision diversification (Rubinstein, 2002):

There is a deck of 100 cards that is composed as follows:

- ❑ 36 cards are green (G)
- ❑ 25 cards are blue (B)
- ❑ 22 cards are yellow (Y)
- ❑ 17 cards are red (R) [original: brown]

5 cards are drawn at random and put into 5 separate envelopes A, B, C, D, E.

Imagine that you receive a price for predicting the correct color of the card in each envelope. What would you predict for the different envelopes.

Results:

42% (Study 1) and 38% (Study 2) of the participants made the *optimal prediction*: Green for each of the 5 envelopes.
The other participants used a diversification strategy.
About 30% of the participants used PM (about 50% of participants who diversified).

Comment:

Please note that even with sampling without replacement, the probability of selecting the green card is greatest for each of the five draws.

5.5.1.2.2 PM and the search for patterns

Various authors argued that PM might indicate peoples' attempt to search for patterns (e.g. Vulkan, 2000; Wolford, Newman, Miller, & Wig, 2004). An experiment of Gaissmaier & Schooler (2008) provides some evidence in favor of this explanation.

In their first experiment they presented participants with two sequences. The first one was a purely random sequence comprising 288 trials with two possible events: A red square (*R*) with p = 2/3 and a green square (*G*) with p = 1/3. The second sequence also comprised 288 trials. However, for this sequence the fixed pattern *RRGRGRRRGGRR* (12 events) was repeated 24 times.

The results revealed that people performing probability matching in the first phase were more likely to detect the pattern in the second phase. Thus, it may be concluded that PM is an indication of peoples' attempts to search for patterns (but see Koehler & James (2008), for a critical comment).

5.5.1.2.3 Individual differences in the usage of PM

Stanovich und West (2000) conducted numerous studies in order to investigate individual differences concerning judgmental errors. With respect to the strategy of probability matching West und Stanovich (2003) observed robust individual differences: People using the optimal strategy exhibited on average a higher cognitive capability (as measured be the SAT [Scholastic Aptitude Test]) compared to persons employing the strategy of probability matching.

Interestingly they also found a distinct difference between sexes: Women tend to use probability matching more frequently than men. For example, in Experiment 3 of West und Stanovich (2003) 41% of the male participants (N = 165) used the optimal strategy whereas 59% resorted to probability matching. For the women (N = 232) the respective rates were 28% and 72%. In each of the three studies sexual differences in the inclination to use probability matching were found, always in the same direction (Total N = 1557). The difference between genders had already been observed in a previous study by Gal and Baron (1996). The reasons for these differences between sexes are unknown.

5.5.1.2.4 PM in animals

Not only humans but also animals exhibit PM. However, as the following example reveals, under specific conditions, animals may exhibit optimal performance whereas humans resort to PM.

Ex. 5-14: Probability matching with rats and students (Gallistel, 1990):

Gallistel (1990, pp. 351-352) reports of a comparison of the performance of rats and students from the year 1960 at Yale University.

A rat was trained to travers a T maze. In 75% the food was located in one branch of the T maze and in 25% in the other one.

Students of a beginner's course in psychology observed the rat and had to predict in each trial which light (invisible for the rat) would turn on indicating the location of food.

Finally, the rat had learned to always turn directly to the branch where the food was predominantly located. By contrast the students exhibited nearly perfect probability matching differing only in about 2 percent from the observed percentages.

Thus, students had to realize that, to their surprise, the behavior of the rat was more intelligent than their own behavior.

Comment:

Under feedback conditions the rat would have exhibited probability matching too (similar to the students) [Cf. Ex. 5-15].

Note:

The presence or absence, respectively of food must not be regarded as proper feedback since the rat did not know whether food had been located in the other branch of the maze.

The strategy of probability has been observed with wild living animals.

Ex. 5-15: Probability matching with animals (Gallistel, 1990):

Smith und Dawkins (1971) studied the hunting behavior of big tits. The birds were allowed to hunt in different places with exact rates of yield. The hunting periods were short enough so that there was no significant reduction of yield in different regions.

One might expect that the birds would (following to an initial exploration phase) remain permanently within the region with the greatest rate of yield. However, this was not the case. Rather *the time of residence of the animals in different regions reflected approximately the relative rate of yield of the region*.

Harper (1982) conducted an experiment with a flock of 31 wild ducks.
Armed with bags containing pieces of bred of 2 g each two experimenters went to the pond on several consecutive days. They positioned themselves 20 meters apart and started to throw the pieces of bred into the water.
The relative rate of throwing of the two experimenters was determined randomly so that it could not be predicted.

Result:

- ❑ At the beginning of the experiment the wild ducks gathered in front of the two throwers where the duration of their residence corresponded approximately to the last rate of throwing.
- ❑ Within 1 minute (during this time about 12 to 18 pieces have been thrown into the water and most ducks had not yet received a piece of bred) the ducks changed their positions in such a way that the relative duration of their stay reflected the actual rate of throwing.
- ❑ In some trials at the end the pieces of bred were thrown with equal rate of throwing. However, the pieces of one thrower were double the size of the other one.
  In this case the ducks first distributed themselves evenly over the two throwers. However after about 5-6 minutes they adjusted their distribution thus reflecting the product of size and rate of throwing.

5.5.1.3 Explanations of Probability Matching

The tendency to probability matching becomes reasonable if the two (or more types) of events are considered as sources with different yield (as is certainly the case for the experiments with the animals). It seems that participants and animals try to get a gain from each of the sources. However, in trying to get a gain from a source with lower yield one neglects the source with the higher yield resulting in a total yield that is below the optimal value. Thus, people using PM do not seem to understand the principle *that the diversification of behavior is suboptimal since reliance on the less yielding source is accompanied by wasting the more abundant source.*

An experiment of Rubinstein (2002) provides evidence for this explanation. The experiment is a variant of the Experiment on Ex. 5-13, presented above:

Ex. 5-16: Weak decision diversification (Rubinstein, 2002):
Imagine you are a detective in a shopping center. Every day a messenger arrives with an envelope.
The center has 4 doors: *G*, *B*, *Y*, and *R*. The doors have different frequency of the messengers entering a door, specifically:

- 36% of the messengers enter door *G*.
- 25% of the messengers enter door *B*.
- 22% of the messengers enter door *Y*.
- 17% of the messengers enter door *R*.

You have to take photos of the messengers as they enter one of the doors for each of the 5 days of the week. However, you have only one camera that you can install each morning at one of the four gates. Which doors would you choose for each of the 5 days?

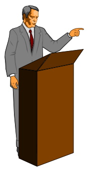
Results:
70% (Study 1) and 72% (Study 2) of the participants always selected door *G* thus exhibiting optimal performance.

This version of the task makes the fact more salient that choosing the less probable alternative amounts to wasting the more probable »source«.
Another explanation that applies to at least some of the findings is based on the representativeness heuristic (cf. Section 4.2.2). According to this explanation, participants assume that the most representative outcome is also the most probable. The most representative outcome is the one that reflects most closely the probabilities of the generating process. For instance, in Ex. 5-13 (p. 278) concerning the envelopes people might think, that the colors of the cards in different envelopes should reflect in some way the given probabilities (see also the example below in Section 0).

5.5.1.4 Non-optimality of Probability Matching

Gallistel (1990) argues that probability matching may be a useful strategy in a competitive situation: If an animal would explore the place with the highest yield only it would attract competitors resulting in a shortage of the resource. In this case, animals exploring locations with lower yield are in a more favorable position. By consequence, the strategy of residing in locations with maximal yield only *is not an evolutionary stable strategy*.
Under the condition of competence the strategy to adjust the duration of residence according the rate of yield seems a more useful behavior

since it enables the animal to gain their share of yield even in the face of numerous and strong competitors.
According to my opinion this line of reasoning is not sound. Even in case of competence it would be more useful to reside at that location that provides the greatest yield where the competence situation has to be taken into account, i.e., the animal has to choose the locations that provide the greatest yield given the competitors at that location.
However, due to bounded rationality these computations may be too complex thus resulting in a simpler strategy the might lead to acceptable results (*satisficing*).



Comment 5-1: Arguments concerning the rationality of probability matching.

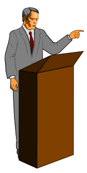
There exist various attempts to justify probability matching and providing arguments that should demonstrate that this strategy is perfectly rational at least under certain conditions. According to my opinion none of these attempts (of which I am aware of) has been successful.

It is also the case that people understand that PM is not the optimal strategy once the principle has been explained to them (Koehler & James, 2008).

5.5.1.5 PM: A Final Test

Consider the following lottery:

> 20 digits are drawn with replacement from the set {0, 1}. The probability of drawing a 0 is p = .9 (thus the probability of drawing a 1 is p = .1). The number of possible lots is 2^{20} = 1'048'576.
> You have the chance to buy a single ticket. Which of the following tickets would you purchase?
>
> (a) The lot with zeros only: 00000 00000 00000 00000.
>
> (b) A lot with 19 zeros and 1 one, e.g. 01000 00000 00000 00000.
>
> (c) A lot with 18 zeros and 2 ones, e.g. 01000 00010 00000 00000.
>
> (d) A lot with 17 zeros and 3 ones, e.g. 01000 00010 00001 00000.



Comment 5-2: Empirical test the probability of the appearance of different patterns

Using the program R, one can immediately test for the appearance of different patterns by executing repeatedly the command:

```
sample(x = c(0,1), size = 20, replace = T, p = c(.9, .1))
```

By the way, one can also simulate quite easily the envelope example (cf. in Ex. 5-13 [p. 278]):

```
x <- rep(c("G", "B", "Y", "R"), c(36, 25, 22, 17))
sample(x = x, size = 5, replace = F)
```

Note that sampling *without* replacement is performed.

5.6 Exercises

Exercise 5-1: Allais paradox

Given:

Problem I:

Option A 1 Mio Sfr. with probability of p = 1.00.

Option B
- 1 Mio Sfr. with probability of p = 0.89.
- 5 Mio Sfr. with probability of p = 0.10.
- 0 Mio Sfr with probability of p = 0.01.

Problem II:

Option A'
- 1 Mio Sfr. with probability p = 0.11.
- 0 Mio Sfr. With probability p = 0.89.

Option B'
- 5 Mio Sfr. with probability of p = 0.10.
- 0 Mio Sfr with probability of p = 0.90.

Result:

Most people choose Option A in *Problem I* Option A and Option B' in *Problem II*.

(a) Demonstrate that peoples' preferences are inconsistent using the Urn representation of Ex. 5-10.

(b) Show that peoples' choices are in contradiction to subjective expected utility (SEU) theory.

Exercise 5-2:

Given:

An experiment comprising three possible outcomes (for each trial): A red lamp lights up in 60%, a green lamp lights up in 30%, and blue lamp lights up in 10% of the cases.

The single trials are independent with the specified probabilities for lighting up.

The task of the participants consists in predicting the color of the lamp in the next trial.

Please determine:

(i) The expected percentage of correct predictions if the participant uses the strategy of probability matching.

(ii) The expected percentage of correct predictions if the participant uses the optimal strategy.

6. References

Abelson, R. P. (1963). Computer simulation of "hot cognition". in S. S. Tomkins & S. Messick (Eds.), *Computer simulation of personality* (pp. 277-302). New York: Wiley.

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35,* 303-314.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369-406.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological Review, 94,* 192-210. http://dx.doi.org/10.1037/0033-295X.94.2.192

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends in Cognitive Science, 8,* 170-177.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Science, 18,* 177-185.

Ayal, S., & Beyth-Marom, R. (2014). The effects of mental steps and compatibility on Bayesian reasoning. *Judgment and Decision Making, 9,* 226-242.

Ayton, P. & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition, 32,* 1369-1378.

Balakrishnan, J. D. (1999). Decision processes in discrimination: Fundamental misrepresentations of signal detection theory. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1189-1206.

Balakrishnan, J. D., & MacDonald, J. A. (2002). Decision criteria do not shift: Reply to Treisman. *Psychonomic Bulletin & Review, 9,* 848-865.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences, 30,* 241–297.

Bargh, J. A. (1994). The four horseman of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (2nd edition, Vol.1, pp. 1-40). Hillsdale, NJ: Erlbaum.

Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer, Jr. (ed.), *Advances in social cognition* (Vol. 10, pp. 1-61). Mahwah, NJ: Erlbaum.

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Bassok, M., & Novick, L. R. (2012). Problem solving. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford library of psychology. The Oxford handbook of thinking and reasoning* (pp. 413-432). New York, NY, US: Oxford University Press. doi: 10.1093/oxfordhb/9780199734689.013.0021

Benedetti, F. (2014). *Placebo effects* (2nd edition). Oxford, UK: Oxford University Press. [HPAED X-2496]

Birnbaum, M. H. (1983). Base rates in Bayesian inference: Signal detection analysis of the cab problem. *American Journal of Psychology, 96*, 85-94.

Böcherer-Linder, K., & Eichler, A. (2019). How to improve performance in Bayesian inference tasks: A comparison of five visualizations. *Frontiers in Psychology, 10,* 267. doi: 10.3389/fpsyg.2019.00267

Brase, G. L. (2002). Ecological and evolutionary validity: Comments on Johnson-Laird, Legrenzi, Girotto, and Caverni's (1999) mental-model theory of extensional reasoning. *Psychological Review, 109, 4*, 722-728. doi: 10.1037//0033-295X.109.4.722

Brase, G. L. (2008). Frequency interpretation of ambiguous statistical information facilitates Bayesian reasoning. *Psychonomic Bulletin & Review, 15,* 284-289. doi: 10.3758/PBR.15.2.284.

Brase, G. L., Cosmides, L., & Tooby, J. (1998). Individuation, counting, and statistical inference: The role of frequency and whole-object representations in judgments under uncertainty. *Journal of Experimental Psychology: General, 123*, 3-21.

Carroll, J. S. (1978). The effect of imaging an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology, 14,* 88-96.

Casscells, W., Schönberger, A. & Grayboys, T. B. (1978). Interpretation by physicians of clinical labratory results. *New England Journal of Medicine, 299,* 999-1001.

Chapman, L. J., & Chapman, E. J. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74*, 193-204.

Cohen, J., Cohen, P., West, S. G, & Aiken, L. S. (2003). *Applied multiple regression/correlation for the behavioral sciences* (3rd edition). Mahwah, N.J.: Erlbaum. [HPAED Pc-14A]

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, *4*, 317-370.

Chernoff, E. J., & Sriraman, B. (2014). *Probabilistic thinking: Presenting plural perspectives.* Dordrecht: Springer. [HPAED PC-490]

Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology, 47,* 738-748.

Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin, 90,* 218-244. .

Cosmides, L & Tooby, J (1992). Cognitive adaptations for social exchange. In: Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). Oxford: Oxford University Press..

Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all ? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58*, 1-73.

Darley, J. M., & Batson, C. D. (1973). “From Jerusalem to Jericho”: A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology, 27,* 100-108.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17-22.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249-267). Cambridge, England: Cambridge University Press.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmutz (Ed.), *Formal representation of human judgment* (pp.17-52). New York: Wiley.

Edwards, W. (1982). Conservativism in human information processing. In D. Kahneman, P. Slovic, P., & A. Tversky. (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 359-369). Cambridge, England: Cambridge University Press.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap.* Boca Raton: Chapman & Hall.

Ernst, E., & Singh, S. (2009). *Gesund ohne Pillen - was kann die Alternativmedizin?* München: Hanser.

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13,* 378-359.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59,* 255-287.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition, 11,* 295-306.

Evans, J. St. B. T., & Stanovich, K. E. (2013a). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*, 223-241. doi: 10.1177/1745691612460685

Evans, J. St. B. T., & Stanovich, K. E. (2013b). Theory and metatheory in the study of dual processing: Reply to comments. *Perspectives on Psychological Science*, *8*, 263-271. doi: 10.1177/1745691613483774

Falk, R. (1992). A closer look at the probabilities of the notorious three prisoners. *Cognition, 43*, 197-223.

Fischoff, B. & Beyth, R. (1975). “I knew it would happen”: Remembered probabilities on once-future things. *Organizational Behavior and Human Performance, 13,* 1-16.

Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science, 66,* 362-378.

Fong, T. G., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*, 253-292.

Fong, T. G., & Nisbett, R.E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology: General, 120,* 34-45.

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology, 65,* 45-55.

Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, *109,* 416-422.

Gal, I. & Baron, J. (1996). Understanding repeated simple choices. *Thinking and Reasoning, 2,* 81-98.

Gallistel, C. R. (1990). *The organization of learning.* Cambridge, Mass.: MIT Press.

Gerry, M, Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imaging a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review, 3,* 208-214.

Gigerenzer, G. (1991). From tools to theory: A heuristic of discovery in cognitive psychology. *Psychological Review, 98,* 254-267.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky (1996). *Psychological Review, 103,* 592-596.

Gigerenzer, G. (2001). Content-blind norms, no norms, or good norms? A reply to Vranas. *Cognition, 81,* 93-103.

Gigerenzer, G, Gaissmaier, W. Kurz-Milcke, E, Schwartz, L. M., & Woloshin, S. (2008). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest, 8,* 53-96.

Gigerenzer, G. Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 513-525.

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684-704.

Gigerenzer, G. & Murray, D. (1987). *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum.

Gigerenzer, G. & Regier, T. (1996). How do we tell an association from a rule? Comment on Sloman (1996). *Psychological Bulletin, 119*, 23-26.

Gilovich, T., & Savitsky (2002). Like goes with like: The role of representativeness in erroneous and pseudo-scientific beliefs. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp.617-624). Cambridge, UK: Cambridge University Press.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*, 295-314.

Girotto, V., & Gonzales, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. *Cognition, 78,* 247-276.

Hacking, I. (2001). *An introduction to probability and inductive logic*. Cambridge: Cambridge University Press.

Hacking, I. (2009). *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference* (2nd edition). London: Cambridge University Press.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgment. *Journal of Experimental Social Psychology, 12,* 392-407.

Harper, D. G. C. (1982). Competitive foraging in mallards: Ideal free ducks. *Animal Behaviour, 30,* 575-584.

Hawthorne, J. (2011). Confirmation theory. In: P. S. Bandyopadhyay, & M. R. Forster (Eds.), *Handbook of the Philosophy of Science (Vol.7): Philosophy of Statistics* (pp. 333-389). Amsterdam: Elsevier.

Herrnstein, R. J., & Murray, C. A. (1996). *The Bell curve: Intelligence and class structure in American life* (2nd edition). New York: Simon & Schuster.

Hertwig, R., & Gigerenzer, G. (1999). The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12,* 275-305.

Hoffrage, U., Krauss, S., Martingon, L., & Gigerenzer, G. (2015). Natural frequencies improve Bayesian reasoning in simple and complex tasks. *Frontiers in Psychology, 6, 1473*. doi: 10.3389/fpsyg.2015.01473

Howson, C., & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach* (3rd edition). Chicago: Open Court.

Ilic, D., Neuberger, M. M., Djulbegovic, M., & Dahm P. (2013). Screening for prostate cancer. *Cochrane Database of Systematic Reviews, 1,* CD004720. doi: 10.1002/14651858.CD004720.pub3.

Janis, I. L., & Mann, L. (1979). *Decision making: A psychological analysis of conflict, choice, and commitment.* New York: Free Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, S. D. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88,* 67-85.

Jungerman, H., Pfister, H. R., & Fischer, K. (2010). *Die Psychologie der Entscheidung* [*The psychology of decision*] (3.Auflage). Heidelberg: Spektrum.

Juslin, P. (2001). Representative Design: Cognitive science from a Brunswikian perspective. In K. R. Hammond, & Steward, T. S. (Eds.), The essential Brunswik: Beginnings, explications, applications (pp. 404-408). Oxford: Oxford University Press.

Kahana, M. J. (2012). *Foundations of human memory*. New York: Oxford University Press.[HPAED Q-2034]

Kahn, H. A., & Sempos, C. T. (1989). *Statistical methods in epidemiology*. New York: Oxford University Press. [HPAED PC-304]

Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp.46-81). New York: Cambridge University Press.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less. Adding a better end. *Psychological Science, 4,* 401-405.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251.

Kahneman, D. & Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263-291,

Keren, G. (2013). A tale of two systems: A scientific advance or a theoretical stone soup? Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science, 8,* 257-262. doi: 10.1177/1745691613483474

Keren, G. & Lewis, C. (1999). On the difficulties underlying Bayesian reasoning: A comment on Gigerenzer and Hoffrage. *Psychological Review, 106,* 411-416.

Keren, G. & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science, 4,* 533-550.

Kleiter, G. (1994). Natural sampling: Rationality without base rates. In G. H. Fischer & D. Laming (Eds.), *Contributions to* *mathematical psychology, psychometrics, and methodology* (pp. 375-388). New York: Springer.

Koehler, D. J., & James, G. (2008). Probability matching in choice under uncertainty: Intuition, versus deliberation. *Cognition, 113*, 123-127.

Krantz, D. (1981). Improvements in human reasoning and an error in L. J. Cohen's. *Behavioral and Brain Sciences, 4,* 340-341.

Krauss, S, & Wang, X. T. (2003). The psychology of the Monty Hall problem: Discovering psychological mechanisms for solving a tenacious brain teaser. *Journal of Experimental Psychology: General, 132,* 3–22. doi: 10.1037/0096-3445.132.1.3

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480-498.

Kruglanski, A. W. (2013). The default interventionist perspective as a unimodel: Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science, 8,* 242-247. doi: 10.1177/1745691613483477

Kruglanski, A. W., & Gigerenzer, G. (2011). Intuitive and deliberate judgments are based on common principles. *Psychological Review, 118,* 97-109.

Larkey, P. D., Smith, R. A., & Kadane, J. B. (1989). It's okay to believe in the "hot hand". *Chance: New directions for Statistics and Computing, 2,* 22-30.

Levi, I. (1981). Should Bayesians sometimes neglect base rates? *Behavioral and Brain Sciences, 4,* 342-343.

Linton, S. J., & Mehlin, L. (1982). The accuracy of remembering chronic pain. *Pain, 13,* 281-285.

Loftus, E. F. (1980). *Memory.* Reading, Mass.: Addison Wesley.

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 585-589.

Lord; F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin, 68,* 304-305. .

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114,* 185-199.

Macdonald, K., Germine, L., Anderson, A., Christodoulou, J., & McGrath, L. M. (2017). Dispelling the myth: Training in education or neuroscience decreases but does not eliminate beliefs in neuromyths. *Frontiers in Psychology, 8,* 1314. doi: 10.3389/fpsyg.2017.01314.

Marcus, G. B. (1968). Stability and change in political attitudes: Observed, recalled and explained. *Political Behavior, 8,* 19-31.

Mather, M., Johnson, M. K., & DeLeonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology, 24,* 437-458.

McCloskey, M., & Zaragossa, M. (1985). Misleading postevent information and memory for events: Arguments against the memory impairment hypothesis. *Journal of Experimental Psychology: General, 114,* 1-16.

McNeil, B. G., Pauker, S. G., Sox, S. C., Jr., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine, 306,* 1259-1262.

Meder, B., & Gigerenzer, G. (2014). Statistical thinking: No one left behind. In: E. J. Chernoff, & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp.127-148). Dordrecht: Springer.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67,* 371-378.

Miller, A. G., Gillan, B., Schenker, C., & Radlove, S. (1973). Perception of obedience to authority. *Proceedings of the 81st Annual Convention of the American Psychological Association, 8,* 213-225.

Mitchell, K. J. & Johnson, M . K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving, & F. I. M. Craig (Eds.), *The Oxford handbook of memory* (Chapter 12, pp. 179-195). Oxford: Oxford University Press.

Mitchell, K. J. & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin, 135,* 638-677.

Morris, M. W., & Peng, K. (1994) Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology,* 67, 949-971.

Mosteller, F. (1965). *Fifty challenging problems in probability with solutions*. Reading, MA: Addison Wesley.

Mueser, R. P., Cowan, N., & Mueser, K. T. (1999). A generalized signal detection model to predict rational variation in base rate use. *Cognition, 69,* 267-312.

Müller, O. (2019). *Altern, Sterben, Tod: Die Vergänglichkeit des Menschen aus der Sicht der Natruwissenschaften [Aging, dying, death: Human perishability from the point of view of the natural sciences]*. Gütersloh: Gütersloher Verlagshaus.

Newstead, S. E., Thompson, V. A., & Handley, S. J. (2002). Generating alternatives: A key component in human reasoning? *Memory & Cognition, 30,* 129-137.

Nickerson, R. S. (1996). Ambiguities and unstated assumptions in probabilistic reasoning. *Psychological Bulletin, 120*, 410-433.

Nickerson, R. S. (2004). *Cognition and chance: The psychology of probabilistic reasoning.* Mahwah, NJ: LEA. [HPAED Q-2314]

Nisbett, R. E., & Borgida, (1981).

Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P.W. (1987). Teaching reasoning. *Science, 238,* 625-631. https://www.jstor.org/stable/1700478

Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of human judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review, 11,* 988-1010.

Osman, M. (2013). A case study: Dual-process theories of higher cognition: Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science, 8,* 248-252. doi: 10.1177/1745691613483475

Paik, M. (1985). A graphical representation of a three-way contingency table: Simpson’s paradox and correlation. *American Statistician, 39,* 53-54.

Pickrell, J. E., Bernstein, D. M., & Loftus, E. F. (2004). Misinformation effect. In R. Pohl (Ed.), *Cognitive Illusions: A handbook of fallacies and biases in thinking, judgment and memory* (pp. 345-361). Hove, UK: Psychology Press.

Pietromonaco, P. R., & Nisbett, R. E. (1982). Swimming upstream against the fundamental attribution error: Subjects’ weak generalizations from the Darley and Batson study. *Social Behavior and Personality, 28,* 641-651.

Pinker, S. (1997). *How the mind works.* New York: Norton.

Pinker, S. (2003). *Das unbeschriebene Blatt: Die moderne Leugnung der menschlichen Natur.* Berlin: Berlin Verlag.

Polman, E. (2010). Why are maximizers less happy than satisficers? Because they maximize positive and negative outcomes. *Journal of Behavioral Decision Making, 23,* 179-190.

Popper, K. R. (1959). The propensity interpretation of probability. *British Journal of the Philosophy of Science, 10*, 25-42.

Popper, K. R. (1984). *Objektive Erkenntnis: Ein evolutionärer Entwurf* (4. Auflage). Hamburg: Hoffmann & Campe.

Redelmeier, D., Katz, J., & Kahneman, D. (2003). Memories of colonoscopy: A randomized trial. *Pain, 104,* 187-194.

Rock, I. (1957). The role of repetition in associative learning. *The American Journal of Psychology, 70*, 186-193.

Rode, C., Cosmides, L., Hell, W., & Tooby, J. (1999). When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. *Cognition, 72*, 269-304.

Roediger, H. L., & Gallo, D. A. (2004). Associative memory illusions. In R. Pohl (Ed.), *Cognitive Illusions: A handbook of fallacies and biases in thinking, judgment and memory* (pp. 308-326). Hove, UK: Psychology Press.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 803-814.

Ross, L, Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control and biases in social-perception processes. *Journal of Personality and Social Psychology, 35,* 485-494.

Ross, L., Lepper, M. R., Strack, F., & Steinmetz, J. L. (1977). Social explanation and social expectation: The effect or real and hypothetical explanation upon subjective likelihood. *Journal of Personality and Social Psychology, 35*, 817-829.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96,* 341-357.

Rotton, J. & Kelly, I. W. (1985). Much ado about the full moon: A meta-analysis of lunar-lunacy research. *Psychological Bulletin, 97,* 286-306.

Rubinstein, A. (2002). Irrational diversification in multiple decision problems. *European Economic Review, 46*, 1369-1378.

Salsburg, D. (2001). *The lady tasting tea: How statistics revolutionized science in the twentieth century.* New York: Freeman,

Schacter, D. L. (1999). *Wir sind Erinnerung: Gedächtnis und Persönlichkeit.* Hamburg: Rowohlt.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1-66.

Schum, D. A. (1994). *The evidential foundations of probabilistic reasoning.* New York: Wiley.

Schustack, M. W. & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General, 110*, 101-120.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130*, 380-400.

Shafer, G., & Vovk, V. (2005). *The origins and legacy of Kolmogorov's Grundbegriffe.* Available from: http://www.probabilityandfinance.com/articles/04.pdf

Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: the mediating effect of ease of imagery. *Personality and Social Psychology Bulletin, 11*, 118-127.

Shiffrin, R. M., & Schneider, W.(1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127-190.

Shimojo, S. & Ichikawa, S. (1989). Intuitive reasoning about probability: Theoretical and experimental analyses of the "problem of three prisoners". *Cognition, 32,* 1-24.

Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the reality of the conjunction fallacy. *Memory & Cognition, 30,* 191-198.

Sinayev A. & Peters, E. (2015) Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6,* 532. doi: 10.3389/fpsyg.2015.00532

Sirota, M., Kostovičová, L., & Vallée-Tourangeau, F. (2015a). Now you Bayes, now you don't: Effects of set-problem and frequency-format mental representations on statistical reasoning. *Psychonomic Bulletin & Review, 22,* 1465-1473. doi: 10.3758/s13423-015-0810-y.

Sirota, M., Kostovičová, L., & Vallée-Tourangeau, F. (2015b). How to train your Bayesian: A problem-representation transfer rather than a format-representation shift explains training effects. *The Quarterly Journal of Experimental Psychology, 68,* 1-9, doi: 10.1080/17470218.2014.972420

Sloman, S. A. (1996a). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3-22.

Slovic, P. Fischoff, B., & Lichtenstein, S. (1982). Facts versus fear: Understanding perceived risk, In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.* Cambridge, England: Cambridge University Press.

Slovic, P. & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgement. *Organizational Behavior and Human Performance, 6,* 649-744.

Slovic, P, Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment, and risk communication: The effects of using actual cases, providing instructions, and employing probability versus frequency formats. *Law and Human Behavior, 24,* 271-296.

Smith, J. M. N., & Dawkins, R. (1971). The hunting behavior of individual Great Tits in relation to variations in their food density. *Animal Behaviour, 19,* 695-706.

Stainton, R. J. (2006). *Contemporary debates in cognitive science.* Malden, Mass: Blackwell. [HPAED Q-1877]

Stanovich, K. E. (2010). *Decision making and rationality in the modern world*. Oxford, UK: Oxford University Press. [Q-1920]

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate?. *Behavioral and Brain Sciences, 23,* 645-726.

Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94,* 672-695.

Talboy, A. N., & Schneider, S. L. (2017). Improving accuracy on Bayesian inference problems using a brief tutorial. *Journal of Behavioral Decision Making, 30,* 373–388. doi: 10.1002/bdm.1949

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4,* 469-477.

Thompson, V. A. (2013). Why it matters: The implications of autonomous processes for dual process theories: Commentary on Evans & Stanovich (2013). *Perspectives on Psychological Science, 8,* 253-256. doi: 10.1177/1745691613483476

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristic-and-biases tasks. *Memory & Cognition, 39,* 1275-1289.

Tu, Y-K., Gunnell, D., Gilthorpe, M. S. (2008). Simpson's paradox, Lord's paradox, and suppression effects are the same phenomenon – the reversal pardox. *Emerging Themes in Epidemiology, 5, 2*.
Available from: http://www.ete-online.com/content/5/1/2.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*, 31-48.

Tversky, A., & Gilovich, T. (1989a). The cold facts about the "hot hand" in basketball. *Chance: New directions for Statistics and Computing, 2,* 16-21.

Tversky, A., & Gilovich, T. (1989a). The "hot hand": Statistical reality or cognitive illusion? *Chance: New directions for Statistics and Computing, 2,* 31-34.

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*, 105-110.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185,* 1124-1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211,* 453-458.

Tversky, A. & Kahneman, D. (1982a). Judgments of and by representativeness. In: D. Kahneman, P. Slovic, & A. Tversky, (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84-98). Cambridge, England: Cambridge University Press.

Tversky, A., & Kahneman, D. (1982b). Causal schemas in judgments under uncertainty. In: D. Kahneman, P. Slovic, & A. Tversky, (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 117-128). Cambridge, England: Cambridge University Press.

Tversky, A. & Kahneman, D. (1982c). Evidential impact of base rates. In: D. Kahneman, P. Slovic, & A. Tversky, (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 153-160). Cambridge, England: Cambridge University Press.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293-315.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representations of uncertainty. *Journal of Risk and Uncertainty, 5,* 297-323.

Van Zandt, T. (2000). ROC Curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26,* 582-600.

von Randow, G. (1994). *Das Ziegenproblem: Denken in Wahrscheinlichkeiten.* Hamburg: Rowohlt.

Vranas, P. B. M. (2000). Gigerenzer's normative critique of Kahneman and Tversky. *Cognition, 76,* 179-193. .

Vranas, P. B. M. (2001). Single-case probabilities and content-neutral norms: A reply to Gigerenzer. *Cognition, 81,* 1059-111.

Vulkan, N. (2000). An economist's perspective on probability-matching. *Journal of Economic Surveys, 14,* 101-118.

Wagenaar, W. A. (1970). Appreciation of conditional probabilities in binary sequences. *Acta Psychologica, 34,* 348-356.

Wainer, H., & Brown, M. L. (2007). Three statistical paradoxes in the interpretation of group differences: Illustrated with medical school admission and licensing data. In: C. R. Rao, & S. Sinharay (Eds.), *Handbook of statistics, Vol.26: Psychometrics* (pp. 893-918). Amsterdam: Elsevier. .

West, R. F., & Stanovich, K. E. (2003). Is probability matching smart? Associations between probabilistic choices and cognitive ability. *Memory & Cognition, 31,* 243-251.

Wilson, A., & Ross, M. (2004). Illusions of change and stability. In R. Pohl (Ed.), *Cognitive Illusions: A handbook of fallacies and biases in thinking, judgment and memory* (pp. 379-396). Hove, UK: Psychology Press.

Wolford, G., Newman, S., Miller, M. B., & Wig, G. (2004). Searching for patterns in random sequences. *Canadian Journal of Experimental Psychology, 58,* 221-228.

Yamagishi, K. (2003). Facilitating normative judgments of conditional probability: Frequency or nested sets? *Experimental Psychology, 50,* 97-106.

7. Books

Baron, J. (2008). *Thinking and deciding.* Cambridge: Cambridge University Press. [FSES M121, TEB-26109, BP2 Economy]

A comprehensive exposition concerning the topics thinking and deciding. Unfortunately, not very exciting to read since the book is a bit verbose.

Dawes, R. M. (1988/2001). *Rational choice in an uncertain world*. Philadelphia: Harcourt Brace Jovanovich. [HAED Q-1160, HPAED Q-1819].

An interesting book on selected topics with respect to deciding and judgments. Competently written the book is quite informative.

Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth.* New York: Free Press. [HPAED Q-1510]

An interesting book that discusses critically the problem of expertise in psychology.

Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press. [SPAED L-2-508]

A collection of articles about human judgments and biases in human judgments. Most of the articles have been published previously in various journals. The book may be conceived of as a continuation of Kahneman, Slovic und Tversky (1982) [see below].

Kahneman, D. (2011). *Thinking fast and slow.* New York: Farrar, Straus, and Giroux.

A recent summary of the most important results concerning judgment and decision biases containing biographical elements concerning the two most important researches in the field [Tversky and Kahneman]. According to my view, the book is a bit too long. Specifically, the two original articles in the appendix appear superfluous.

Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press. [HPAED Q- 508]

A collection of articles concerning human judgments and biases in human judgments. Informative but a little bit boring.

Lilienfeld, S. O., Lynn, S. J., Lohr, J. M. (2003). *Science and pseudoscience in clinical psychology.* New York: Guilford Press. HPAED X-1615.

In this book various experts discuss problems, controversies and questionable practices in psychotherapy.

Lilienfeld, S. O., Lynn, S. J., Rusco, J., & Beyerstein, B. L. (2010). *Fifty great myth of popular psychology: Shattering widespread misconceptions about human behavior.* Chichester: Wiley-Blackwell, [HAED Q-1927]

An informative and mostly amusing presentation of popular erroneous theories in psychology.

Nisbett, R. E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of human judgment*. Englewood Cliffs, NJ: Prentice-Hall. [HPAED Q-1060]

This book is a classic and one of my hits. Despite its age the book is still very informative and readable. It is very well written and amusing. A special feature is the fine self-irony of the authors.

Plous, S. (1993). *The psychology of judgment and decision making*. New York: McGraw-Hill. [HPAED Q-1157]

A compact, informative, and well written summary of the most important aspects concerning human decisions and judgments. The book has received the William James Book Award.

Pohl, R. (2004). *Cognitive Illusions: A handbook of fallacies and biases in thinking, judgment and memory.* Hove, UK: Psychology Press. [HPAED Q-1713]

The book presents an overview of the psychology of judgmental errors. According to my opinion, the rationality debate presented in the book is not very conclusive. In addition, the book contains numerous errors.

Rosenzweig, P. (2007). *The halo effect and eight other business delusions that deceive managers.* New York. Free Press. [HPAED W-2077]

The book presents negative consequences of ignoring the Halo effect in (half-) scientific investigations concerning successes and failures of corporations. In addition, various other judgmental errors afflicting the assessment of managers and corporations are explicated.
The book is very entertaining, easy to read and really interesting.

Stanovich, K. E. (2010). *Decision making and rationality in the modern world.* Oxford, UK: Oxford University Press. [HPAED Q-1920]

Concise, informative and well written. The book contains a good treatment of the rationality debate.

Sutherland, S. (2007). *Irrationality.* London: Pinter & Martin. [HPAED Q-1900]

Amusing, well written, the book provides an overview of a broad palette of judgmental errors; easy to read.