



A Tribute to Kahneman and Tversky in the Context of Mathematics Education

Ulrich Hoffrage¹ and Laura Martignon²(✉)

¹ Faculty of Business and Economics (HEC Lausanne), University of Lausanne, Lausanne, Switzerland

² Institute of Mathematics, Ludwigsburg University of Education, Ludwigsburg, Germany
martignon@ph-ludwigsburg.de

Abstract. Daniel Kahneman and Amos Tversky have changed the way we think about human rationality. Based on their contributions, the Enlightenment notion of rational humans thinking according to the laws of logic and probability had to be replaced by a more realistic conceptualisation of human thinking. However, their experiments on the so-called “errors” of the human mind were countered by a plethora of other experiments aimed at identifying the conditions that reduce or even eliminate these errors. Both the discoveries of Kahneman and Tversky, and the reactions they provoked, also had an impact on the way we teach critical thinking and probability in school. Maths educators have been prompted by these discoveries to find new, more adaptable ways to address reasoning and decision making under uncertainty in the classroom. This paper presents tools and methods that have been developed to promote logical reasoning in tasks on the so-called conjunction rule and tasks on strictly probabilistic reasoning. These tools use adequate and dynamic representations partly inspired by Barbara Tversky’s work on dynamic visualizations and the way the human mind constructs thoughts. The two studies reported here were carried out in schools and demonstrate the effectiveness of these tools.

Keywords: Conjunction Fallacy · Intensional · Extensional · Icon Arrays

1 Introduction

A famous minister from the German region of Bavaria has repeatedly insisted on giving a good reason for forbidding marijuana. According to his reasoning, marijuana is a gateway drug to Cocaine. He has cited that more than 70% of cocaine users have their drug initiation with marijuana. Without pretending to defend marijuana, we do claim that there is a problem with his reasoning. What matters here is not the conditional probability that someone who makes use of cocaine today began by using marijuana in the past; the relevant conditional probability is that someone who starts with marijuana ends up using cocaine. This number is much smaller.

There are endless examples of wrong conditioning in reasoning by politicians, but also influencers and marketing agents. Recent advertisement posters in French subway

stations suggest that young men should try to drive like women do, because 88% of young people's mortal car accidents are caused by young men. Here again what matters for the assertion is the reverse conditional probability, namely the proportion of young men who have severe accidents, compared to that of women and adjusted for driven kilometer. This proportion could easily be well below one per thousand, and if men drove more than 88% of the kilometers, they would drive even more safely than women. This is just one of the many flaws human reasoning can exhibit, as has been empirically demonstrated in the famous collection of human fallacies in reasoning – not just probabilistic – discovered by Kahneman and Tversky and their close colleagues during the second half of the last century. According to the discoveries they made (see, for instance, Eddy 1982), people are bad at estimating the predictive value of a feature, like the result of a test for detecting a disease. People even make mistakes when assessing the probability of a conjunction of events, committing the so-called *conjunction fallacy*. Understanding how and why all those so-called flaws come about is a first step towards eliminating them and towards critical thinking.

2 How Can School Foster Critical Thinking?

The OECD has declared Critical Thinking as one of the relevant competencies for the 21st century. School is the main institution that can foster critical thinking of all individuals in modern times, as education should be granted every citizen of a civilized society. Education stands, in more than one sense, as opposed to Nudges. While nudges only change features of the decision-making environment, education boosts citizens' conscious judgment and decision-making capabilities. Education should therefore be more robust when it comes to initiating behavioral change across different environments (Hertwig and Grüne-Yanoff 2017).

To make it more visible and tangible to practitioners, the OECD has worked with networks of schools and teachers in 11 countries to develop and test a set of pedagogical resources that exemplify what it means to teach, learn and make progress in critical thinking in primary and secondary education. Our intervention programs have been inspired by the OECD practices. Our intention has also been to implement experience-based approaches based on activities supported by digital plugins. We begin by describing our intervention studies on the conjunction problem.

3 Extension Versus Intension in the Conjunction Fallacy

The original, by now classical, paper by Tversky and Kahneman (1983) on the conjunction fallacy was entitled “Extensional vs. Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment”. From our perspective, deeply influenced by the “Logique de Port Royal”, that paper illustrated the importance of the intensional vs. extensional dichotomy in the context of reasoning and decision (see Stenning et al. 2017, for a detailed treatment). Let us briefly recall the meaning of the terms “intensional” and “extensional”: whereas the intensional meaning (or the intension of a term) refers to the qualities or attributes that the term connotes, the extensional meaning (or the extension of a term) consists of the list of all items or individuals the term denotes. The pair of terms

originated from the environment of Aristotelian logic and was established as ‘epitome of the term’ (“comprehension de l’idée”) and ‘scope of the term’ (“étendue de l’idée”) by the Port-Royal’s logic. In Mathematics, the distinction between intensional meaning and extensional meaning was of great importance when, at the end of the 19th century, the fundamentals of the discipline were written down by scientists like Russell or by Frege. A set can namely be described intensionally or extensionally. Take, for instance, the set of even numbers between 1 and 10, formally written as $\{x \mid x \text{ is a natural number divisible by 2, larger than 1 and less than or equal to 10}\}$, but it can also be written as $\{2, 4, 6, 8, 10\}$. The first formal expression describes the set intensionally, while the second expression is purely extensional. To put it somewhat casually, intensional descriptions of meaning are composed of attributes (or predicates) of the described item, while extensional descriptions are lists of the objects with those attributes.

Going back to the conjunction error, we want to stress that it was conceived extensionally: the probability of a conjunction should be rated as smaller than that of each of its components because the set that corresponds to the conjunction is contained in each of the sets of its components (for the first full and strictly extensional treatment of the fallacy, see Hertwig 1994).

Let us remember what the conjunction fallacy is about. Tversky and Kahneman (1983) demonstrated this fallacy with a short vignette about Linda:

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.

Subjects are then given some statements about Linda that they should rank according to their probability of being true, among them:

- Linda is a bank teller
- Linda is a bank teller and a feminist.

The Linda sketch naturally and strongly evokes an intensional interpretation as a story, enunciating qualities of Linda (i.e., a thumbnail biography). However, when ranking the probabilities that Linda is a bank teller, or that she is both a bank teller and a feminist, a switch to an extensional interpretation is triggered. Thus, while a detective searching for Linda among bank tellers would certainly look for a woman with political interests – and why not even a feminist – the probabilist knows that there are more bank tellers in the world than bank tellers who are also feminists. The detective thinks intensionally, while the probabilist measures sets extensionally.

Conjunctions and their probabilities are usually taught in schools based on Venn diagrams, which are ovals representing sets. This representation expresses sizes of sets but not of populations. We posit that an even more extensional approach based on representations like icon arrays foster the understanding of probabilities to an even higher degree than working with Venn diagrams. In the next section we describe how an extensional approach can be fostered by creating conjunctions in icon arrays.

4 Creating Your Own Conjunctions

An experience-based approach to learning about extensional aspects of conjunctions can be implemented by playing with the following dynamical website created by the second author and Tim Erickson: <https://www.eeps.com/projects/wwg/wwg-en.html>. The third section on that website is called the Conjunction Playground. Here the user can, in a playful manner, create her/his conjunctions in icon arrays. The user can choose between two task contents, namely, Soccer and Chess (one with 25 and another one with 100 icons), or Computers and Gardening (with 36 icons) and will subsequently see the icon array displayed in Fig. 1 (left panel) or that of Fig. 2 (left panel), respectively. The user plays by dragging the labels he/she finds at the bottom of the icon array, like a football or a chess-token for the icon array depicted in Fig. 1, and a plant or a computer for the icon array depicted in Fig. 2. After the user has played with the content, the icon array may look like the ones displayed in the right panels of Fig. 1 and Fig. 2, respectively.



Fig. 1. The left panel shows the screen after the user had clicked on Conjunction Playground and chose Soccer and Chess as task content. Having then selected icons from the top menu, 25 boys are displayed. At the bottom left side, the user can tip on the soccer-ball icon and drag it on those boys who, in his or her imagination, are good at soccer. In addition, the knight can be dragged on all those boys who are considered to be good at chess. After having dragged the labels on the individual icons, the array may look like that displayed in the right panel.



Fig. 2. For the task content Computers and Gardening, the icon array displays 36 icons of girls. Like for the content displayed in Fig. 1, the user can decide which girls do gardening and which play computer games. The left panel shows the starting point, the right panel a possible result produced by a given user.

After the user has played around in these icon arrays, clicking on icons and dragging features on them, the user can go to the menu at the top of the array. He/she can now click

on a button labeled tree, and on buttons labeled Story 1 and Story 2. To start with the former, the button tree leads to a double tree like those illustrated in Fig. 3. These double trees exhibit the different four conjunction nodes in the middle row: in the situation shown in Fig. 3 (left panel), there two boys who are good both at soccer and at chess, 11 boys who are good at soccer but not at chess, etc. In the right panel of Fig. 3 we see the conjunction nodes in the middle row indicating that 5 girls play computer games and do gardening, another 4 who love gardening but do not play computer games, etc. Note that the numbers in the double tree count all possible combinations of the features (and their margins), while the numbers that are displayed below the icon array (see Figs. 1 and 2) only include some of them (albeit the crucial ones). Both the numbers below the icon arrays (Figs. 1 and 2), and the numbers in the double tree (Fig. 3) are automatically updated with every drag-and-drop of any of the features.



Fig. 3. Double trees that are displayed when the user clicks on trees after having generated the icon arrays shown in the right panels of Fig. 1 and Fig. 2.

Next to icons and trees, the top menu also offers the possibility to open story 1 and story 2. While Story 1 is strictly extensional, Story 2 is intensional. In the case of Soccer and Chess they read as follows:

Story 1 (extensional): Mikkel loves Sudoku and mathematics. He does lots of puzzles and has memorized pi to 50 places. Imagine 100 boys like Mikkel. Are there more boys who are good at soccer, or more boys that are good at soccer and chess?

Story 2 (intensional): Mikkel loves Sudoku and mathematics. He does lots of puzzles and has memorized pi to 50 places. What would you bet: That Mikkel is good at soccer, or that Mikkel is good at soccer and chess?

In the case of Computers and Gardening the two cover stories are:

Story 1 (extensional): Mia is a vegetarian. She eats organic yogurt and is the president of the Ecology Club. Imagine 36 girls like Mia. Are there more girls who play computer games, or more girls that play computer games and do gardening?

Story 2 (intensional): Mia is a vegetarian. She eats organic yogurt and is the president of the Ecology Club. What would you bet: That Mia plays computer games, or that Mia plays computer games and does gardening?

The intervention study we implemented (Study 1) required from students, each one operating from his/her laptop or ipad, that they freely created their own conjunctions enactively. They tended to place the icons for soccer and chess such that many boys appear to be good at soccer while only a few appear to be good at chess. An example for this kind of assignment is illustrated in Fig. 1, where 13 boys are good at soccer but only 6 are good at chess. The conjunction of being good at soccer and chess consists of 2 icons only. Students can easily see that there are more boys who are good at chess than boys who are good at both. Note that this insight conflicts with thinking along Kahneman and Tversky's representativeness heuristic. This heuristic suggests that (1) being good at soccer is typical for boys and that, hence, (2) the conjunction (which explicitly contains the term soccer) is more likely than the single event being good at chess (which does not contain the typical event).

We propose that the experience gained by playing around on that website leads to an embodied understanding of extensional conjunctions. Stated as a hypothesis:

Hypothesis: Gaining experience by creating conjunctions in icon arrays will foster extensional reasoning, thereby protecting from committing the conjunction fallacy.

To test whether this hypothesis can be empirically supported, we have implemented two intervention studies to demonstrate the effectiveness of such an experience-based approach.

5 Study 1: Testing the Facilitating Effect of Enacting Visual Representations in Conjunction Tasks

Both intervention studies were based on comparisons between two school classes (one with 28 and the other with 26 school students as participants) who were put in the treatment condition and two school classes (one with 26 and the other with 27 school students as participants) who served as a control condition. Our participants were 15 years old and in the 9th class of a German Gymnasium. All classes had the same pre-test with questions about probability and one conjunction task with the cover story about Mikkel in Story 2. The treatment consisted of two school hours (each 45 min) of experience-based work with the conjunction playground concluded by diagrams on the black board under the title Conjunction of events and a conceptual explanation. The control classes had 2 h of exercises on solving equations. Four weeks after the intervention treatment, all four classes were subject to the same post-test. The post-test contained two tasks that corresponded to the Linda task both in the cover story and in the question to be answered. This means that the cover stories were intensional descriptions, like in Story 2 both in the case of Mikkel and of Mia. The wordings were different from those in the Mikkel and Mia tasks, but referred again to persons or animals, with two different predicates.

5.1 Results of Pre- and Post-tests

The results are displayed in Table 1. There was practically no difference between the two conditions in the pre-test. Almost all students committed the conjunction fallacy, as predicted and repeatedly shown by Kahneman and Tversky. In the treatment condition, only 11.1% of the students (6 of 54) made correct extensional choices for conjunctions in both conjunction tasks the pre-test contained, and in the control condition, only 1.9% (1 of 53) made correct extensional choices. The pattern changed, however, in the post-test. In the treatment group, after having accumulated experience on the website, 85.2% of the students (46 of 54) provided correct choices for the two conjunctions in the post-test. Not surprisingly, there was no such improvement in the control condition: students' error rate stayed on the same high level as it was for the pre-test. We could, hence, observe a significant effect of the experience-based training.

Table 1. Numbers (and proportions) of students who provided correct answers (i.e., extensional choices) in the pre- and post-tests in the treatment and control condition in Study 1 (on conjunction tasks).

	Treatment condition	Control condition
Pre-test	6 (of 54)	1 (of 53)
Post-test	46 (of 54)	1 (of 53)

Following the teachings of Gerd Gigerenzer and his school, who have criticized the use of statistical rituals in cases like this, where the results speak for themselves, to the naked eye, we do not report results of inferential statistic tests here (and also not in Study 2 below). In this case, in which the studies were conducted in one school without randomizations, Gigerenzer would even speak of “mindless statistics” (see Gigerenzer 1993 and 2004).

The website where we find the Conjunction Playground used in our intervention studies offers another scenario, namely that of conditioning an event upon another event, with the usual components. These components are again an icon array and buttons leading to double trees for representation and two stories, a fully extensional one and an intensional one. The conjunction of two events is the basic element for measuring the dependence of one event upon the other. If the event *S* represents, for instance, *is good at soccer* in the small population of boys we treated above, and *C* represents the event *is good at chess*, then the conditional probability of a boy being good at chess given that he is also good at soccer is measured by the proportion of those who are good at both among those who are good at soccer. The two events are called independent if this proportion is the same as the proportion of boys who are good at chess in the whole population. As we described in our introduction, learning to deal with conditional probabilities is fundamental, and is often not easy for humans. To help students and adults overcoming such problems, we have designed interventions that foster conditional reasoning.

Under similar conditions as those described in Sect. 4, we performed another intervention study, Study 2, this time on probabilistic conditioning and Bayesian inference.

Participants were again pupils of the same four 9th classes we had for Study 1 described above. Study 2 was conducted three weeks after Study 1. Those youngsters who were in the treatment group for the conjunction tasks in Study 1 were also the ones for the treatment group for the intervention on probabilistic conditioning and Bayesian inferences in Study 2.

Instead of working with a pre-test we could base our assessment of students' knowledge and understanding of Bayesian reasoning on a test given by the teachers after having completed their teaching on probabilities. The test contained one task based on the following contingency table for an Elisa test for HIV in a certain population of 101.000 people (Table 2).

Table 2. Statistical properties of the Elisa Test, that were used in the HIV task in the pre-test of Study 2

Elisa Test	Disease present	Disease absent	Total
positive	995	500	1.495
negative	5	99.500	99.505
Total	1.000	100.000	101.000

The questions in the task were devoted to finding probabilities of patients satisfying different conditions:

- That a patient with the disease tests positive
- That a patient without the disease tests negative
- That a patient with the disease tests negative
- That a patient without the disease tests positive
- That a patient who tests positive does have the disease

Our intervention in Study 2 consisted of two hours of experience-based learning working with the dynamic page <https://www.eeps.com/projects/wwg/wwg-en.html> (specifically, its first section dubbed “The explanatory power of features”). One of the contents in this first section is a disease in a certain population (with a high base rate), with the icon array displayed in Fig. 4.

Going from Fig. 4, left panel, to Fig. 4, right panel, visualizes the effect of *sorting*, which facilitates the counting. Students experience this action as the first “statistical” move, or even “the beginning of statistics”.

The next statistical step is to add classification to sorting. Classification is based on forming classes according to features. Handy representations can be tables, or, if a hierarchy is desired, trees and double trees, as in Fig. 5.

The tool and the visualizations behind “The explanatory power of features” is designed to support children and adults becoming informed and competent when:

- understanding base rates (of having the disease and of not having the disease),
- dealing both with the sensitivity of a test (i.e., the probability of testing positive when having the disease) or its specificity (i.e., the probability of testing negative when not having the disease),

- inferring the positive predictive value of a test (PPV, i.e., the probability of having the disease when testing positive) and its negative predictive value (NPV, i.e., the probability of having the disease when testing negative),
- understanding how the PPV and the NPV hinges on sensitivity, specificity, and base rates.

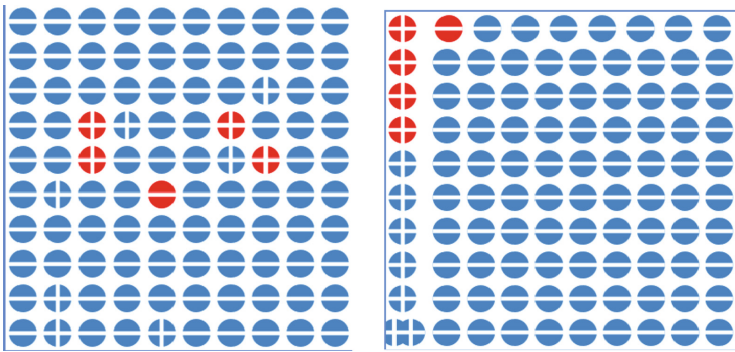


Fig. 4. An icon array for 100 patients who may or not have HIV (left panel: unsorted, right panel: sorted). Legend:  means “disease is absent and test is negative”,  means “disease is absent and test is positive”,  means “disease is present and test is negative”, and  means “disease is present and test is positive”.

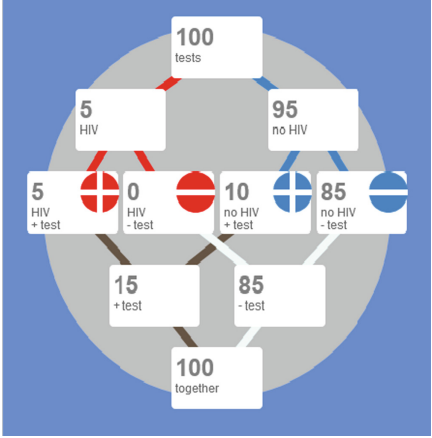


Fig. 5. The double tree arranges counts of the margins of the features “HIV” and “not HIV”, those of the margins of the features “positive test” and “negative test”, and, in the middle row, counts of the conjunctions of these features. The counts in Fig. 5 correspond to those in the icon array displayed in Fig. 4. The double tree facilitates Bayesian reasoning.

The treatment group in our Study 2 had only two school hours of experience-based work with this webpage and a brief conceptualization. The control group had a class on solving equations in correspondence with the schedule. Three weeks later both groups were tested with two tasks with other cover stories on medical tests but equivalent situations and questions. Prior to our intervention, that is, in the pre-test, the two conditions did not differ (Table 3 shows the number of students who answered questions (a) or (e) (or both) correctly, and Table 4 shows the number of students who answered all questions correctly). In contrast, there was a big difference between the two conditions in the post-test, after treatment: While the results in the post-test were excellent for the treatment group (45 school students answered all questions correctly; Table 3), and there were 51 out of 53 who answered either (a) or (e) (or both) correctly (Table 4), there was practically no improvement in the control group, both when considering only questions (a) and (e) (Table 3) and when considering all five test questions (Table 4).

Table 3. Results of pre-tests and post-tests in the treatment and control condition in Study 2 (on probabilistic conditioning and Bayesian inference). The numbers denote how many students answered either question (a) or (e) (or both) correctly

	Treatment condition Questions a) or e)	Control condition Questions a) or e)
Pre-test	21 (of 54)	21 (of 53)
Post-test	52 (of 54)	23 (of 53)

Table 4. Results of pre-test and post-tests in the treatment and control condition in Study 2 (on probabilistic conditioning and Bayesian inference). The numbers denote how many students answered all questions (a) - (e) correctly

	Treatment condition	Control condition
Pre-test	20 (of 54)	17 (of 53)
Post-test	45 (of 54)	23 (of 53)

Our results clearly show that, at least for these students, using this dynamical website with the icon arrays, sorted and unsorted, and the corresponding double trees is a successful method to train extensional reasoning both for assessing conjunctions (Study 1) and for conditional reasoning (Study 2).

6 Conclusion

Daniel Kahneman and Amos Tversky (and their close colleagues) introduced and used extremely interesting tasks that school students can work on and should be able to cope with from 9th class on in order to acquire the competency of critical thinking as a basis for sound decision making. For some of these tasks, for instance, Bayesian inference

tasks, Gerd Gigerenzer and his group developed instruments that could improve adults' performance (like the so-called "natural frequencies"; Gigerenzer and Hoffrage 1995; Hoffrage et al. 2000). Likewise, Gigerenzer and his group introduced icon arrays to communicate medical information to adults (Galesic et al. 2009). Brase (2008) analyzed different representations for conditional and Bayesian reasoning as well, and his results also confirmed the advantages of using such icon arrays. Icon arrays are an excellent representation of extensional information. Our work consists in making use of these ideas in the context of math education, adding the dynamical tools for experience-based learning. We have also been inspired by the work of Gage and Spiegelhalter (2016), who, on their turn, introduced natural frequencies for teaching probability in English schools.

The webpages presented here were inspired by the studies on the advantages of icon arrays as representation of information in probabilistic reasoning tasks mentioned above, adding the component of dynamical, interactive and embodied treatment. Our main contribution is the enactive character of these webpages, which allow for an experience-based learning for probabilistic reasoning (see Martignon and Hoffrage 2019) and a movement based thinking, as Barbara Tversky has consistently proposed (Tversky 2019).

Schools in Germany are slowly beginning to implement such practices and it has been our purpose to implement and test such procedures in order to exhibit their didactical power.

Disclosure of Interests. We declare that we have no competing interest of any kind, related to the work presented here.

References

- Brase, G.L.: The power of representation and interpretation: doubling statistical reasoning performance with icons and frequentist interpretations of ambiguous numbers. *J. Cogn. Psychol.* **26**(1), 81–97 (2014)
- Eddy, D.: Probabilistic reasoning in clinical medicine: problems and opportunities. In: Tversky, A., Kahneman, D., Slovic, P. (eds.) *Judgment Under Uncertainty: Heuristics and Biases*, pp. 249–267. Cambridge University Press, Cambridge (1982)
- Frege, G.: *Die Grundlagen der Arithmetik: Eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Verlag von Wilhelm Koenner, Breslau (1960). (In English)
- Frege, G.: *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number*. Translated by Austin, J. L., 2nd edn. Northwestern University Press, Evanston (1884)
- Gage, J., Spiegelhalter, D.: *Teaching Probability*. Cambridge University Press, Cambridge (2016)
- Galesic, M., Garcia-Retamero, R., Gigerenzer, G.: Using icon arrays to communicate medical risks to low-numeracy people. *Health Psychol.* **28**, 210–216 (2009)
- Garcia-Retamero, R., Galesic, M., Gigerenzer, G.: Do icon arrays help reduce denominator neglect? *Med. Decis. Mak.* **30**, 672–684 (2010). <http://mdm.sagepub.com/content/30/6/672>
- Gigerenzer, G.: The superego, the ego, and the Id in statistical reasoning. In: Keren, G., Lewis, C. (eds.), *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*, pp. 313–339. Erlbaum, Hillsdale (1993)
- Gigerenzer, G.: Mindless statistics. *J. Socio-Econ.* **33**, 587–606 (2004)
- Gigerenzer, G., Hoffrage, U.: How to improve Bayesian reasoning without instruction: frequency formats. *Psychol. Rev.* **102**, 684–704 (1995)

- Hertwig, R.: Why dr. gould's homunculus doesn't like to think like dr. gould: the conjunction fallacy reconsidered. Dissertation at the University of Konstanz (1994)
- Hertwig, R., Grüne-Yanoff, T.: Nudging and boosting: steering or empowering good decisions. *Perspect. Psychol. Sci.* **12**(6), 973–986 (2017)
- Hoffrage, U., Lindsey, S., Hertwig, R., Gigerenzer, G.: Communicating statistical information. *Sci.* **290**, 2261–2262 (2000)
- Martignon, L., Hoffrage, U.: Wer wagt, gewinnt? Wie Sie die Risikokompetenz von Kindern und Jugendlichen fördern können. Hogrefe, Göttingen (2019)
- Russell, B.: *The Principles of Mathematics*. Cambridge University Press, Cambridge (1903)
- Stenning, K., Martignon, L., Varga, A.: Probability-free judgment: integrating fast and frugal heuristics with a logic of interpretation. *Decision* **4**(3), 136–158 (2017)
- Tversky, A., Kahneman, D.: Extensional vs. intuitive reasoning: the conjunction fallacy in probability judgment. *Psychol. Rev.* **90**, 293–315 (1983)
- Tversky, B.: *Mind in Motion: How Action Shapes Thought*. Basic Books, New York (2019)