Connecting Basic Proportional Thinking with Reasoning About Risks

Ulrich Hoffrage[1], Laura Martignon[2(✉)], Tim Erickson[3], and Joachim Engel[2]

[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

[3] Eeps Media, Oakland, CA, USA

Abstract. Risk literacy requires basic elements of numeracy. It requires some ease with basics of probability theory. Yet, it is these basics which often cause difficulties, in particular when they are presented using abstract formalisms. This paper reviews a systematic framework for representations of information that eliminate difficulties and frequent fallacies in dealing with probabilities. A large subfamily of these representations is inspired by Otto Neurath's isotypes and consists of so-called icon arrays. Another subfamily contains trees, which are hierarchical noncyclic graphs. Yet another subfamily consists of double trees that foster intuitions for Bayesian inferences. An interactive webpage is presented that can be used by both adults and children, with buttons and sliders to set parameters and with 3 different levels of statistical literacy. Furthermore, trees are examined as structures for combining multiple cues in order to classify situations under risk and make decisions. Plugins for constructing such trees and for reckoning with risks are presented and discussed.

1 Introduction

Humans have dealt with risks since the beginnings of human history. Yet the rigorous, formal treatment of risks is a modern achievement. This mathematical treatment is essentially based on probability theory and was cemented during the early nineteenth century with the work of De Morgan on probability and life contingencies (1838). It attained full formal rigor only during the first decades of the last century, being embedded in the edifice of Mathematics. This edifice, solidly built on axioms and theorems proven by means of inferences based on classical logic, can appear daunting to untrained lay people. In fact, precisely this formal rigor hinders the natural approach to „doing" mathematics based on fruitful intuitions. The tension between rigorous, formal mathematics and mathematical intuitions remains a hot topic of mathematics education, especially at school level: *what* should be taught in school and *how* should it be taught so that school students acquire mathematical competencies beyond procedural techniques? The tension between the axiomatic treatment of Kolmogorov and the intuitive, quasi- empirical treatment of Pascal and Laplace, based mainly on numerical proportions, is so strong that it led Leo Breiman (1968) to state that "probability theory is condemned to having a right

J. Baratgin et al. (Eds.): HAR 2023, LNCS 14522, pp. 386–406, 2024.
https://doi.org/10.1007/978-3-031-55245-8_25

and a left hand—the right hand being the measure-theoretical approach that guarantees mathematical rigor, and the left hand meaning 'intuitive probabilistic thinking'" (p. 7). The modern concept of a probability is defined by means of a real-valued function on a sigma algebra of subsets of a set that satisfies certain axioms. The enthusiasm for this definition in the first decades of the 20th century was enormous. It proved, once again, that set theory had become, as David Hilbert put it, "A paradise, from which no-one will throw us out." (Hilbert 1926).

In the sixties, mathematicians both in Europe and in the United States prompted the introduction of set theory in schools. This enthusiasm was enhanced by the possibility of representing sets by means of Venn diagrams. This representation of sets was at hand since John Venn had introduced his diagrams in the nineteenth century (Venn 1880). Venn diagrams represent sets, their intersections and their unions by circles or ovals that may overlap or be nested within each other. In the sixties and seventies, these Venn diagrams were introduced as representations of sets in primary and secondary schools in most European countries and in several other countries around the world. To the dismay of mathematicians, school students, their parents and even many of their teachers were irritated and frustrated, to say the least. We cite here a typical example: Stephanie Krug, who went to school in Baden Württemberg back in the seventies and whom we interviewed, recalls her reluctance to draw triangles and circles in different colors placed in those Venn diagrams (Fig. 1). "What for?" she asked.

Fig. 1. From a notebook of Stephanie Krug in 1974

In fact, the reaction to set theory and Venn diagrams in schools represents one of those rare victories of teachers united with parents in many countries of the world. In Germany, the protests from teachers and parents were so strong that set theory was banned from primary school, and with it the Venn diagrams that had confused everyone.

Convinced that probability theory requires both set theory and the functions defined on sets, German mathematics educators were reluctant to introduce probabilities earlier than in advanced secondary school. This tendency changed somewhat during the nineties, and elements of probability theory and statistics were covered in short chapters in school math books. Sadly, these chapters were often left to be treated at the end of school years, with little emphasis and little enthusiasm. In this article, we focus on that realm of elementary stochastics, which examines the possible relationships between two bivariate variables and is essentially based on proportional reasoning. We posit that proportional reasoning, supported by dynamic visualizations of "natural frequencies" (see below for a definition of this term), can provide effective intuitions of probabilistic situations and risk literacy. We exhibit digital plugins for working with proportions related to risks, that have successfully been used at different institutions.

2 Proportions, Frequencies, and Probabilities

With the results of cognitive psychologists during the second half of the twentieth century on people's dealing with probabilistic inferences, things became even worse for the acceptance of probabilities in math education. Experiments by psychologists such as Amos Tversky and Daniel Kahneman discredited Breiman's left (intuitive) hand. Using many examples, Kahneman and Tversky (e.g., 1974) had shown that people can have enormous difficulties in dealing with various tasks that involved probability judgments. Note that in these tasks the necessary information was typically expressed in terms of probabilities. Likewise, they had shown that people have a hard time to correctly compare the probability of compound events with the probabilities of the constituents of these compounds. In the opinion of these two psychologists and many others who followed their line of research, people are unable to handle probabilities. More generally, the conclusion of these authors' heuristics-and-biases program was that people are not "rational".

This pessimistic view of human thinking did not go unchallenged. It provoked new empirical work as well as theoretical and methodological discussions, which were driven, to a considerable extent, by Gerd Gigerenzer and his students. In many experiments, these authors have demonstrated that so-called cognitive illusions can be made to disappear (e.g., Gigerenzer 1991, Gigerenzer, Hertwig, Hoffrage, & Sedlmeier 2008). They have argued that sometimes the "wrong" statistical (or logical) norm has been applied or that the stimulus materials used in experiments have not been representative of participants' natural environment to which they have adapted. Most important for the present article were, however, their argument and their demonstrations that *information needs representation* and that performance in judgment tasks can improve tremendously when information is presented in terms of frequencies instead of probabilities. Some of the studies by Gigerenzer and his students then led to a complete redesign of Venn diagrams, which made a big difference for children and adults. Ovals with small abstract figures such as triangles and squares are not helpful. What helps, however, are grids with representations of individuals, items of all sorts, mythical creatures, or animals that are easy to sort and count. It is easy to choose content that is appealing and motivating even for young children. With these representations, elementary probabilistic thinking and risk literacy boils down to reckoning with proportions, comparing them, and drawing conclusions from comparisons.

3 Classifications in Risky Situations

The main scope of this work is to address how elementary proportions are at the basis of risk literacy. Let us recall that one of the main theorems of probability theorem states that in any aleatory experiment which can be repeated „ad infinitum", the relative frequencies converge to the real probability. This result basically implies that relative frequencies, which are proportions of successful results divided by the total number of results, are the concrete, palpable approximations of probabilities. Mathematics educators often insist that working well with these approximations is a sufficient basis for risk literacy. Decisions in risky situations often depend on classifications of situations, which, on their turn, depend on the features or cues that characterize them. For instance, a doctor classifies a patient as “in high risk of heart attack” based on certain features extracted from the electrocardiogram of the patient and from behavioral cues, like the intake of certain medicines or chest pain. Once the patient is classified as “high risk” he/she is sent to the coronary care unit. If not, that is if he/she is not “at high risk” then the patient can be assigned a regular nursery bed. Medical situations are one of the great application fields in decision making, where risk literacy becomes fundamental.

Observe that in most medical situations, one feature alone is not enough for making a good decision. The immense progress in medicine and epidemiology is precisely the discovery of tests, symptoms, and behavioral traits that can fully characterize a patient’s risk situation, so that accurate decisions tailored to the specific risk can be made.

The Covid pandemic that recently shocked the world is a wellknown example of medical decision making under high risk. Which are the relevant cues and how can their reliability be measured? We had some direct experience because we, at least one of the authors, had to deal with the disease in the region of the German city of Tübingen. It was interesting to consult doctors who worked together with the main Hospital of Tübingen and capture their simple strategy for decision making in case of symptomatic patients. Basically, the decision tree of doctors during the months of April and May 2020 looked like this:

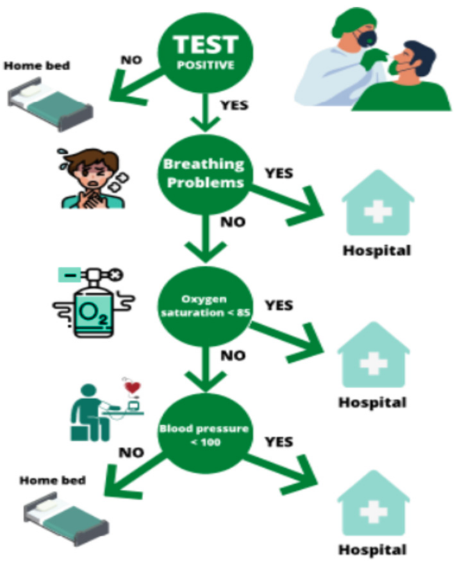

Fig. 2. A frequently adopted classification and decision tree for symptomatic patients at the beginning of the Covid pandemic

Observe that a positive test did not lead to immediate hospitalization. Breathing problems led to hospitalization, but, in its absence, other cues were checked. The order of those other cues was important here. In general, research has shown that a robust and accurate tree similar to the one shown in Fig. 2 can be constructed by ordering the cues according to their validity, that is, the proportion of correct classifications. However, the validity of the cues considered at the beginning of the Covid pandemic was not yet established by large empirical studies, so that the situation was characterized by uncertainty rather than risk (Mousavi & Gigerenzer 2014). Nevertheless, there were some preliminary numbers and estimates that could be used to construct such trees and so we witnessed the development of these decision trees as an ad hoc process, in the sense that "the science had to be developed along the way" without a large base of prior knowledge.

4 Scaffolding Risk Literacy

Being able to construct tools, like trees, for decisions under risk, is one component of risk literacy. Which are other components? In what follows, we briefly present our (Martignon & Hoffrage 2019) four-stage model of risk literacy. It consists of four components (Fig. 3):

I) Detecting risk and uncertainty
II) Analyzing and representing uncertain or risky situations
III) Comparing alternatives and dealing with trade-offs
IV) Making decisions and acting

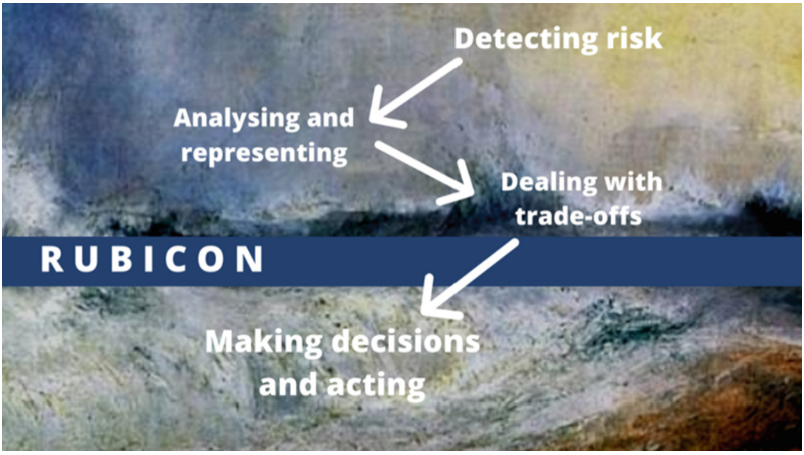

Fig. 3. Scaffolding Risk Literacy by means of four components (adapted from Martignon & Hoffrage 2019).

Of these components the first one – detecting and identifying risks and uncertainties in ordinary life – requires sufficient psychological disposition, either innate or acquired during childhood and youth. The second component – analysing and modelling risks – is not only psychological but, above all, adaptive. It requires basic skills but also basic education, especially in these times when we are confronted with a sheer amount of quantitative information on a daily basis – information that requires and provokes the acquisition of basic numeracy skills. The availability of information goes hand in hand with the availability of digital media and its tools, including those that can improve risk literacy. Concepts of risk and related tools need to be understood and trained in order to improve basic numeracy skills and risk literacy. To illustrate: if a person's Covid-19 test was negative and she was told that the false negative rate of the test was, say, 2%: is she now at risk? And what if the test was positive? Answers to this type of question became crucial to citizens all over the world during 2020. But it is, of course, not just the Covid 19 pandemic which prompts our need to understand the validity of features and how it can be computed. For instance, when a woman is told that regular screening reduces the risk of breast cancer by 50%, what should she do? Or if we are told that eating bacon sandwiches increases our risk of getting bowel cancer by 20%, how seriously should we try to avoid bacon sandwiches? One message of this paper is that there are simple tools and principles for analysing and modelling risks and uncertainties, so that these become amenable to being assessed and compared in terms of elementary proportions that can provide the basis for sound decisions, even of young students. The third component of risk literacy – namely comparing alternatives and dealing with trade-offs – builds on the first two, but goes beyond them and therefore represents an additional skill. This paper is concerned with the second, third and fourth components of risk literacy.

5 Proportional Thinking and Logical Principles at the Foundation of Classification

A child of 10 years can deal with the following situation: Consider 25 pupils in a classroom, some are boys, and some are girls. Some have short hair. Some are boys and have short hair. Some wear skirts, some wear trousers. Some wear skirts and have short hair. Would one bet that a child of our class who wears a skirt is a girl? Most probably we would. But would we bet that a child with short hair is a boy? Probably not! Features or characteristics, like „short hair" or „wears a skirt", are the essence of classification and inference in everyday life and we should early learn to deal with them. We extract features, items, and concepts out of situations with ease, we are able to classify based on features and tend to define situations, items, and concepts based on features.

In this example, three bi-variate variables were mentioned: Gender (boy/girl), Hair length (short/long), and Dress (trousers/skirt). The variables (here, three) and the individual objects (here, 25 pupils) can be conceived of as two poles. To start with the former: the variables can be used to define classes and to classify objects. They are, hence, on a population level and define what all have in common that belong to this class. Note that 25 is, in the present example, the entire population; but from a higher point of view this number can be seen as describing just one sample of a population that is, in principle, unlimited – and in fact, there are other classrooms of the same school, other schools in the same city, other cities in other countries, and other cohorts in coming years. The direction from "class to object", or from "population to individual" is deductive.

Conversely, one could also start from the other pole, the individuals. To make it concrete: A teacher could ask her 25 pupils to put all tables aside and to gather all in the middle of the room. She could then send some to the left side of the room – Peter, Thomas… and Rafael go here, and Sandra, Yvonne …, and Kim go there – and let the pupils detect her organizing principle. While the teacher proceeds in a deductive manner and applies a general rule that determines whether a given individual goes left or right, the pupils whose task it is to detect the rule, need to make an inductive inference: They need to ask what does the one group of individuals have in common, what does the other have in common, and what discriminates between them? Having detected this principle, they should be able to predict where the teacher would place a new pupil (that is, to make an out-of-sample prediction if the door would open and pupil number 26 enters).

This little exercise, which can easily be implemented in the classroom, illustrates the difference between deductive reasoning (from population to sample, or top-down) and inductive reasoning (from sample to population, or bottom-up). When the teacher uses a variable to form groups, she uses deductive reasoning ("All girls should be on the left side – you are a girl – therefore you go left"). When she does not reveal her criterion but asks pupils to find it out, they engage in inductive reasoning. Note that this distinction is akin to Piaget's (1956) distinction between intensional and extensional reasoning about features in a given sample, which goes back to the Port Royal Logic. Features are intensional aspects of elements of sets, like „wearing a skirt". This is a variable that characterizes an unlimited number of objects and specifies what the members of this category (or set, or subpopulation) have in common. In contrast, if we start with a list of the names of all children wearing a skirt in our class, then we are performing an extensional operation rendering all children wearing skirts by listing them. The question whether extension and intension can be treated in one framework goes back a long way: The dichotomy can be found at the heart of what is considered the second epoch of logic initiated by Antoine Arnauld and Pierre Nicole in their book "Logic or the Art of Thinking" which was published in 1662. The dichotomies we just discussed (deductive vs. inductive; top-down vs. bottom-up, intensional vs. extensional) are intimately related to the two visual representations that we already mentioned in our introduction. Let us consider the deductive, top-down, and intensional viewpoint first. It is directional: from population to sample. It goes from features that characterize an, in principle, unlimited (and hence, not countable) population to (countable) individuals that possess these features. In other words, it goes from qualities (variables) to quantities (individuals). This viewpoint focusses on sets, but not just on how many individuals are in these sets. *Structure* (classes defined by features) comes first – *content* (individuals as carriers of features) second. A way to visualize sets and structure is to use adequate "good-old" Venn-diagrams. The regions or areas depicted in these diagrams specify what *all* have in common who are in a specific area, but these individuals (who, together, build the "all" in a specific area) are not individually identified.

The other viewpoint – inductive, bottom-up, and extensional – is also directional, but now from individual to population. The starting point is constituted by "countable individuals", say pupils in a classroom. By inspecting them closer and by comparing them, the question arises how they can be described, what they have in common, and, eventually, how they can be distinguished from each other, that is, which variables could be used to describe and to classify them. A way to visualize countable individuals is to use icons for representing each of them. A given icon depicts features of a given individual, but these features (along which individuals can be distinguished) are not displayed as a set.

Venn-diagrams and icons can be seen as the two polar representations that visualize the starting points of the two perspectives described above. Venn-diagrams visualize sets and icons visualize individuals. Sets as a starting point allow one to use a variable to classify an individual, and individuals as a starting point invites one to ask how they can be described and grouped. To illustrate, the abovementioned teacher in the class with 25 pupils starts with a variable, say gender that defines the set of boys and girls, then looks at her pupils, one after the other, and applies this variable as a classifier to determine, in a deductive manner, who should go to which side of the room. Conversely, those children who find themselves on one side of the room, without having been told anything about the organizing principle, first look at their ingroup and on those on the other side. They hence start from the individuals and then, in a second step, consider potential variables to test whether they could explain the assortment. Finding potential variables to scrutinize them requires inductive reasoning, testing whether a candidate variable can explain the observed grouping requires deductive reasoning.

For each of the two directions described above, we have seen that the natural next step was to leave the viewpoint (as defined by one's standpoint) and go in the direction of the other pole. The deductive view would hence apply intensional reasoning, starting with a classifier, and then look at a given individual in order classify it. This amounts to placing a certain individual into one of the areas in a Venn-diagram. Content is used to fill structure in a top-down manner. Conversely, the inductive view wonders how individuals, represented by icons, could be sorted. By shifting scattered icons around and grouping them according to defining characteristics, sets emerge in a bottom-up manner. Both directions from the two starting points meet in the same middle. No matter whether sets are filled with individuals, or whether individuals are identified as members of sets and sorted accordingly, at the end of each process there is structure with content, or, conversely, content with structure.

The discussion above can be supported by the panels displayed in Fig. 4. Panels A and F of Fig. 4 artificially separate objects and variables. Panel A focuses on objects but is mute about their features. Panel F focusses on features and possible relationships among them but does not contain any countable objects. Panels A and F can thus be seen as two poles: Objects without features and variables without objects, respectively. We already said that in our daily perceptions, these two poles are not separated, and it is hence only straightforward to explore the middle-ground between the two poles. Moving from Panel A towards Panel F leads us to Panel B which adds features to the objects. These features allow for sorting objects, which goes together with grouping them into classes or sets. The result of such grouping is shown in Panel C. Evidently, grouping facilitates counting. Note that sorting, classifying, and counting are elementary statistical operations even children at a very young age are capable of.

Starting at the other pole and moving from Panel F towards Panel A leads us to Panel E which fills the space of possibilities with objects. Compared to Panel F, Panel E leaves the world of pure structure and reminds one that the space consists of countable units – even if, in contrast to the present Panel E, these units are not yet counted, and even if these numbers may be infinite (e.g., repeated outcomes of a chance device like a roulette wheel).

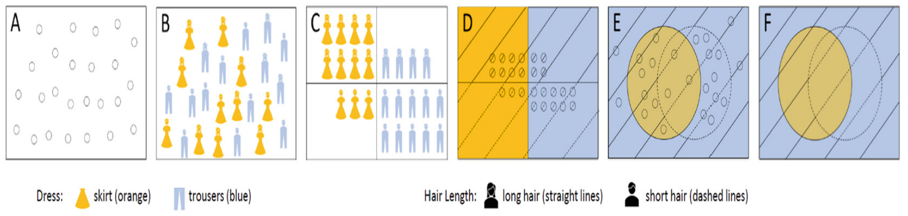

Fig. 4. Panel A: 25 distinguishable objects without any features. Panel B: The same objects, but now described with respect to two variables, dress and hair length. Panel C: The same objects, but now sorted according the two variables. Panel D: The same objects and assortment are now visualized on a more abstract level. Panel E: Visualization of possibilities that arise from combining two dichotomous variables, hair length and dress. For each of the possibilities, the number of objects from a given sample is visible. Panel F: Venn-diagram visualizing the possibilities that arise from combining two dichotomous variables.

Johnson-Laird's mental model theory is essentially based on exactly this step from Panel F to Panel E. When solving reasoning tasks that could be supported by Venn diagrams (e.g., "All P are Q" and "Some Q are R"; Is it true that "some R are not-P"?), mental model theory posits that people do not operate in an abstract variable space, but construct individual instances, thereby searching for examples that confirm or disconfirm the conclusion that they are asked to scrutinize. Arranging the objects of a finite sample that facilitates counting may lead to a representation such as the one depicted in Panel D.

Panels C and D hit the middle ground between the two poles. On the one hand (when coming from Panel A), description and structure are added to otherwise indistinguishable individual objects, and on the other hand (when coming from Panel F), an abstract structure and a space of possibilities is filled with concrete and countable cases. While Panel C maintains the analogous representation of individual cases that were already displayed in Panel B, Panel D inherits the level of abstraction that comes with a focus on possible features and their combinations displayed at the right end of the Figure.

The results on icon arrays, as presented in the next section, clearly indicate that the extensional approach fosters probabilistic intuitions of untrained people. We recall here that it was Otto Neurath who, during the first half of the twentieth century, used and introduced formats such as those depicted in Panel C – he called those little icons *isotypes,* henceforth we will refer to arrays of such isotypes also as *icon grids* or *icon arrays*.

6 Icon Arrays for Risk Literacy: Following Otto Neurath

Icon arrays are a form of graphical representation that illustrates principles for the design of risk communications, inspired by Neurath's Isotypes (Trevena et al. 2013). An icon array is a form of pictograph or graphical representation that uses grids of matchstick figures, faces or other symbols to represent statistical information. An indicator of good quality risk communication is an adherence to the principle of *transparency*, which is definitely a characteristic of icon arrays: such representations define an appropriate reference class, and risks are presented in absolute rather than relative numbers (e.g., 1 out of 1,000 fewer women die from breast cancer with mammography screening as opposed to communicating the relative risk reduction of 20%). Icon arrays can be designed to communicate a variety of statistics transparently, including simple and conditional event frequencies (e.g., conditional probabilities). In medical risk communication, for instance, icons typically represent individuals who are affected by a risk, side-effect, or other outcome. Icon arrays are helpful for communicating risk information because they draw on people's natural tendency to count (Dehaene 1996), while also facilitating the visual comparison of proportions. For example, to represent a 3% risk of infection, icon arrays represent the proportion of individuals who end up with an infection, for instance, an icon array may just depict 100 icons, of which 3 are marked as „special". The one-to-one match between individual and icon has been proposed to invite identification with the individuals represented in the graphic to a greater extent than other graphical formats. Icon arrays are suitable for facilitating the understanding of risk information due to two characteristics: First, they arrange the icons systematically (Fig. 4C) rather than randomly (as in Fig. 4B). Second, and relatedly, they visualize a part-whole relationship.

Institutions devoted to fostering the intuitions of "responsible patients" like the Harding Center in Berlin are interested in propagating basics of risk literacy. Another institution, the AOK, which is one of the main insurance companies in Germany, communicates information to patients by means of fact boxes, as illustrated in Fig. 5.

What does the patient perceive here? The array on the left shows icons for 1,000 women who do not perform screening. Five of them die of breast cancer. The array on the right side shows 1,000 women who undergo screening regularly. Four of these women die of breast cancer, implying that the absolute risk reduction through regular screening is 1 per 1,000 (the reduction from 5 per 1,000 to 4 per 1,000 is 1 per 1,000). The so-called relative risk reduction in the breast cancer situation is 20% for women performing screening (1 per 1,000 whose lives could be saved compared to 5 per 1,000 who would lose their lives without screening; see Sect. 10 for dynamical representations of risk changes).

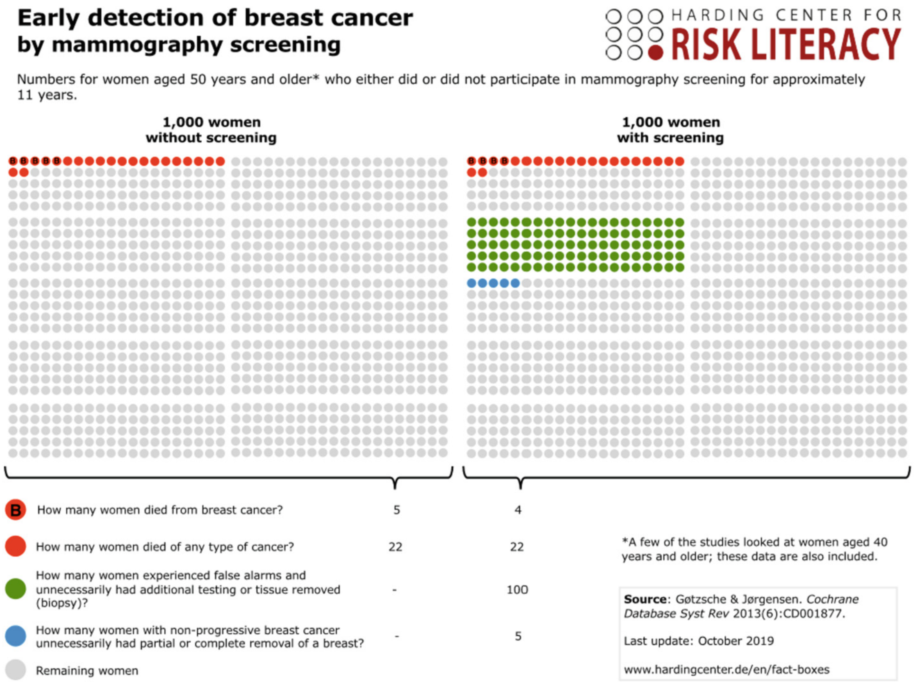

Fig. 5. A fact box designed by the Harding Center and used by the largest health insurance company in Germany, the AOK, for communicating information on the risk reduction caused by regular screening.

7 Beyond Neurath: Dynamic Icon Arrays

Even though icon arrays are already very helpful in facilitating understanding (Garcia-Retamero & Hoffrage 2013), there is still room for improvement: one can make them dynamic. In a dynamic webpage designed by Tim Erickson (https://www.eeps.com/projects/wwg/wwg-en.html), icon arrays can be sorted and organized by a simple click on a button (see, for instance, the HIV example, which is also briefly described below), so that relevant features can be quantified at a glance. ***Sorting*** is the first elementary statistical action we perform, sometimes just mentally (Martignon & Hoffrage 2019). Dynamic displays can thus become particularly useful for communicating about co-occurring or conditional events, which is relevant for understanding the meaning of features in the medical domain. For this purpose, sorting icon arrays becomes essential, as we illustrate with the example of 100 people who were tested as to whether they are HIV positive (Fig. 6).

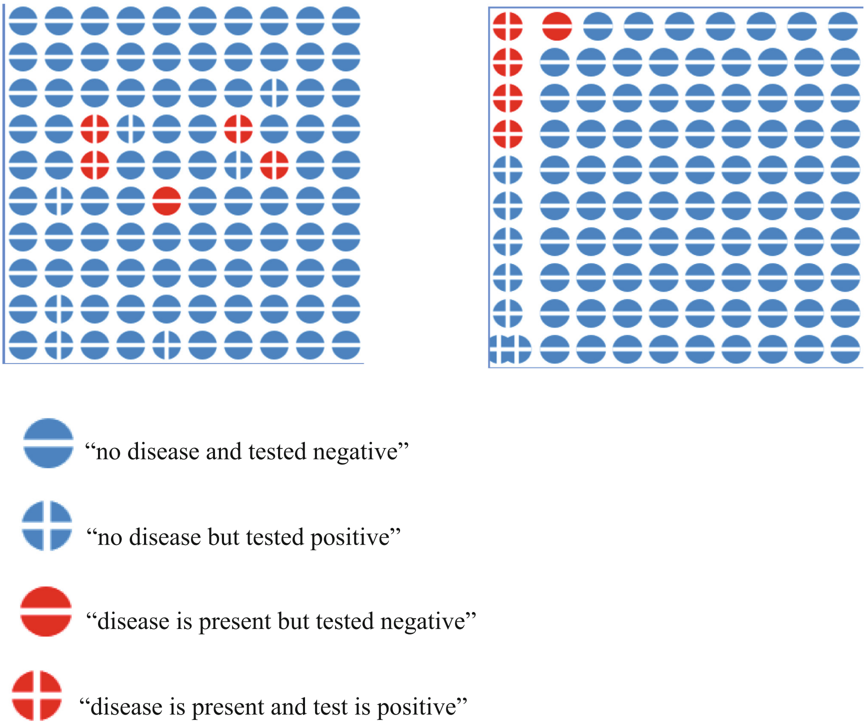

Fig. 6. This icon array, unsorted (left) and sorted (right), represents 100 people, diseased or not diseased, who are tested as to whether or not they are HIV positive.

8 Constructing Trees and Double Trees Starting from Icon Arrays

The validity or predictive value of a feature (e.g., a positive test, or breathing problems) can be computed by means of Bayes' Theorem. Consider, a physician receives new evidence (E) in form of a positive result. To infer whether a certain disease (D) is present or not, the physician should use her prior probability that the disease (D) is present, as well as the two likelihoods of a positive test result (if the disease is present and if the disease is not present, respectively) to calculate the so-called *posterior* probability that the disease is indeed present given the evidence, i.e., the positive test result. The corresponding formula is called *Bayes' Rule*, and was first formulated by the mathematician, philosopher and minister Thomas Bayes in the eighteenth century. Using Bayes' rule, the probability that the disease (D) is present once a new piece of evidence becomes known is calculated as follows:

$$P(D|E) = \frac{P(E|D)P(D)}{P(E|D)P(D) + P(E|\overline{D})P(\overline{D})}$$

Formula 1: Bayes' rule

The formula shows how to solve this evidential reasoning problem: In the medical setting, P(D|E) is the probability that the patient has the disease given that they tested positive on the test. People are notoriously bad at manipulating probabilities, as a plethora of empirical studies have shown (Eddy 1982; Gigerenzer & Hoffrage 1995).

The tree in Fig. 7 represents information about disease and test in a causal, sequential and hierarchical setting by means of a tree: The presence of the disease, represented by D+, is the label on the left node or leaf of the tree in the first level, while D- represents its absence. In this particular case the number 0.01 on the branch between the initial node and the node labeled with D+ in Fig. 7 represents the probability of the disease being present, also called its base rate.

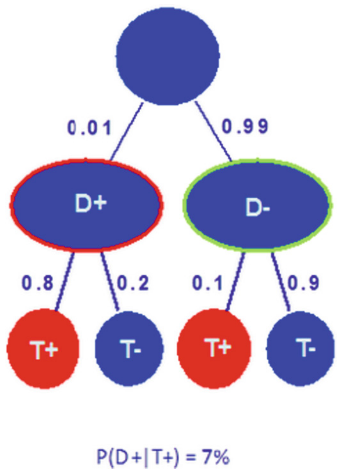

Fig. 7. A tree representing the binary categories "Disease" and "Test result" (D+ means that the disease is present, while D- denotes absence of the disease; T+ and T- denote a positive and negative test result, respectively).

Gigerenzer and Hoffrage (1995) have proposed a didactical simplification or reduction of the initial probabilistic treatment, which deserves the name of a heuristic: the systematic use of so-called natural frequencies. They argued that the kind of reasoning needed to make assessments on the diagnosticity of a test or symptom (or, in general, of a feature characterizing a category) can be facilitated by changing the format of information representation. In the same article (Gigerenzer & Hoffrage 1995) have empirically shown that diagnostic assessments based on new evidence could be substantially improved when the statistical information was provided by means of natural frequencies compared to representation in terms of probabilities.

9 Dynamic Trees of Natural Frequencies

Natural frequencies are the frequencies that naturally result if a sample is taken from a population. In case of one hypothesis or cause, like a disease, and of a piece of binary evidence, like the positive or negative result of a test, natural frequencies are the result of counting members of a given sample in each category. Translating probabilities into natural frequencies is always possible and becomes an ecologically rational heuristic that facilitates reasoning. In medicine, physicians' diagnostic inferences have been shown to improve considerably when natural frequencies are used instead of probabilities (Gigerenzer 1996; Hoffrage, Lindsey, Hertwig, & Gigerenzer 2000). Figure 8 illustrates both approaches to the causal tree: one by means of probabilities and one by means of natural frequencies.

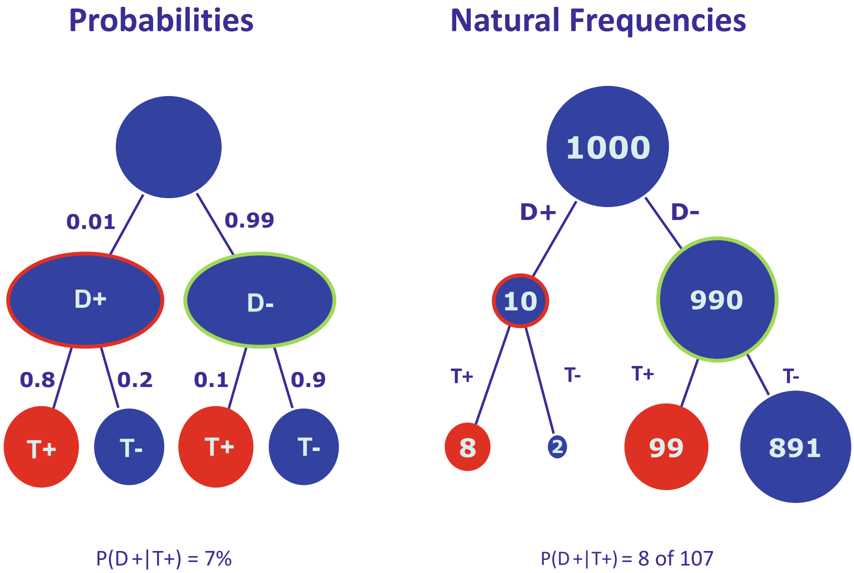

Fig. 8. Two trees, one labeled with probabilities, the other labeled with natural frequencies, representing the knowledge of the physician on a certain patient concerning breast cancer.

Trees with natural frequencies as labels can be constructed in the causal direction, i.e., from cause to evidence, as in the Figures above. The beneficial effect of natural frequencies could also be used as a basis to design tutorials that teach students to better cope with probability representations. Instead of teaching them how to plug probabilities into Bayes' rule, they have been taught how to translate these probabilities into natural frequencies and subsequently derive the solution for there. In an online-tutorial, Sedlmeier & Gigerenzer 2001) could show that such frequency-tree representations were superior to probability trainings (long-term performance of over 90% compared to 20%, respectively). Likewise, representation training also proved to be superior over rule training in a classical classroom setting with medical students, using medical problems as content (Kurzenhäuser & Hoffrage 2002).

Natural frequencies go hand in hand with icon arrays as illustrated by the dynamic web page the reader can reach by means of the QR Code below (https://www.eeps.com/projects/wwg/wwg-en.html)

This page can help the public to become "informed" and "competent" when dealing both with the sensitivity or specificity of a test and with its positive/negative predictive value.

These resources are designed to support instruction of children and adults to become informed and competent when

- dealing both with the sensitivity or specificity of a test,
- dealing with positive/negative predictive values of tests and dependence on base rates,
- understanding base rates,
- understanding relative and absolute risks, and
- understanding the subtleties of features' conjunctions.

To summarise, the resource is designed to make the teaching and training of risk literacy easy and transparent, by offering multiple complementary and interactive perspectives on the interplay between key parameters. Such interactive displays for adults have been introduced, for instance, by Garcia-Retamero, Okan, & Cokely (2012). Clicking on any of the three sections leads to pages where a variety of contexts are presented. For instance, in *The explanatory power of features* one can choose between contexts – one is *Pets and bells*, which is quite appropriate for children of fourth class – and see a display like the following (Fig. 9):

Fig. 9. Representations of 10 pets, cats and dogs, random and sorted.

The natural question is: "If a pet is wearing a bell, is it likely to be a cat?" The task is to judge the validity, or predictive value of this feature for the category *Cats*. The button "group bells together", at the left side under the picture, sorts the pets, so that it becomes easy to visualize pets wearing a bell.

The role of base rates is illustrated through the use of the sliders placed under the array. Maintaining the total number of pets equal to 10 one can enhance the base rate of "cats", while keeping the sensitivity of "bell" constant, as illustrated in Fig. 10.

Fig. 10. In this display the number of pets remains 10 but the base rate of cats is now "6 out of 10"

Wearing a bell now becomes moderately predictive for the category "cats". The next instructional step is to construct trees. A button at the top left of the grid in Fig. 10 leads to the corresponding double tree, illustrated in Fig. 11.

The double tree in Fig. 11 exhibits two inference directions: one is causal the other is diagnostic. The double tree is a simple and transparent way of approaching Bayesian reasoning. Studies by Wassner (2004) clearly demonstrated the effectiveness of such double trees for fostering successful Bayesian reasoning in the classroom. He worked with ninth grade students in Germany.

Fig. 11. The double tree from the icon array in Fig. 10.

10 Other Elements of Risk Literacy by Means of Dynamical Representations

The Webpage "Worth the risk?", also illustrates the subtleties connected with risk reductions and increases in transparent ways that are easy to grasp. Figure 12 shows 20 boys who have had a bike accident, ten of which were wearing a helmet. The faces with a pad and a black eye represent boys with severe injury caused by the bike accident.

Fig. 12. Icon array exhibiting 20 boys having a bike accident, ten of them wearing helmets

Simply sorting the icon array by grouping helmets together allows an easy grasp of the risk reduction provided by helmets.

While all the dynamical resources described so far are devoted to the second component of Risk Literacy, as described in Fig. 3, the third author of this paper has also designed and produced plugins for the third and fourth component. Opening https://codap.xyz/ the reader finds, among many plugins, also ARBOR. This plugin is designed for the construction of trees for classification and decision, based on data concerning features for classification. Decision trees constructed by means of clever algorithms which can be used in complex medical situations exist and have been developed by eminent statisticians, such as CART (Breiman 1996); implementations of those algorithms are available in common software platforms such as R (Erickson & Engel 2023). The trees proposed in ARBOR are utterly simple and intuitive and young students can easily understand them. The basic idea is that the display starts with an initial node that specifies a response variable—the binary criterion or category the tree is designed to predict. Then the user can drag any attribute to any node in order to make (or replace) a branch based on that attribute. Finally, to make a prediction (in the case of a classification tree), the user has to add a "diagnosis leaf" to the end of every branch to indicate what conclusion you should come to if a case arrives at that branch. Let's see what that looks like in practice using a famous dataset about heart patients from Green and Mehr (1999). It has 89 cases with four attributes: MI (whether the patient had a myocardial infarction, a heart attack); pain (whether the patient complained of chest pain); STelev (whether the "ST" segment on an EKG was elevated); and oneOf (whether the patient showed any of four other symptoms). First, the user must set up a *response variable*. This involves deciding which attribute the tree will predict, and which value the tree will orient towards. In our situation, we are trying to predict whether the patient will have a heart attack (MI) using the other three attributes. We also need to orient our tree: will we look at the proportion of patients that do get a heart attack (MI = yes) or the proportion that do not (MI = no)? In Arbor, the response variable and its orientation appear in the "root" node of the tree, represented by a box as shown in Fig. 13.

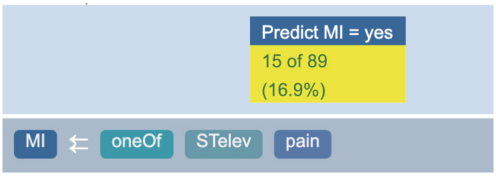

Fig. 13. The positive value of the criterion (MI = Infarction)

Growing the Tree. Now we want to "grow" our tree. We will do so by dragging attributes (oneOf, STelev, or pain) from the palette and dropping them on a node. CODAP's built-in graphing helps us explore the data before constructing a tree. In the case of this question, we can make a graph that shows how the value of MI is associated with pain. The figure shows that, indeed, a patient is more likely to have an MI if they complain of chest pain—but to a student, the relationship probably does not look as they expected. Even with chest pain, a large majority of patients do not get a heart attack (Fig. 14).

The plugin allows then to construct individual trees for each cue and then compose trees with three cues like the one shown in Fig. 15.

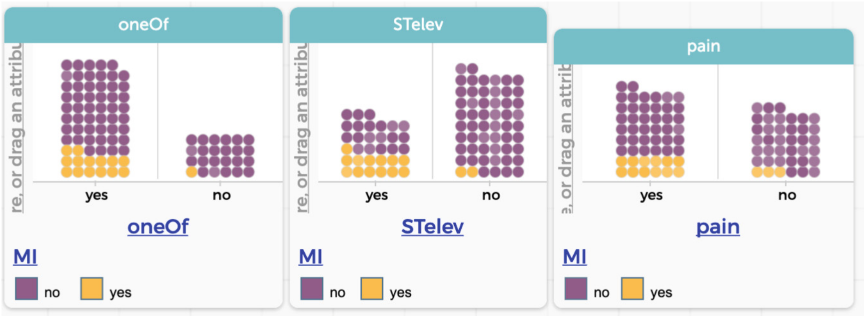

Fig. 14. User-friendly display created by the plugin for the statistics of three cues

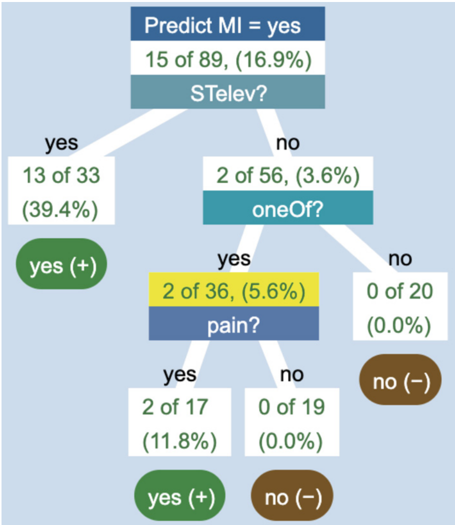

Fig. 15. This tree has been constructed: One can drag and place cues one under the other and construct trees.

This tree could be constructed also with other orderings and we would be able to compare their performances and choose the „best" one. The plugin is thus a facilitating tool for doing the necessary steps for good classification in a simple user-friendly way. There are many other useful plugins in the webpage https://codap.xyz. Another plugin is, for instance, Lotti, which is specifically constructed for component 4 of risk literacy. Here the plugin presents the user with „doors" and alternatives with different outcomes. The plugin provides the typical „lottery" situation, with fixed gain versus situations of risk versus benefits (Fig. 16).

Concluding, we simply stress the benefits of the dynamical, interactive tools for classifications based on dynamic, extensional representations of information, fostering intuitions on risk and provide tools of risk literacy.

Fig. 16. A game for acquiring an understanding of risk through clicking either on plan A, with a fixed allowance, or B with a varying allowance, with larger expected value.

References

Breiman, L.: Probability. SIAM, New York (1968)

De Morgan, A.: An Essay on Probabilities, and Their Application to Life Contingencies and Insurance Offices. Longmans, London (1838)

Dehaene, S.: The organization of brain activations in number comparison: event-related potentials and the additive-factors method. J. Cogn. Neurosci. **8**(1), 47–68 (1996). https://doi.org/10.1162/jocn.1996.8.1.47

Eddy, D.: Probabilistic reasoning in clinical medicine: problems and opportunities. In: Slovic, P., Tversky, A. (eds.) Judgment under Uncertainty: Heuristics and Biases, pp. 249–267. Cambridge University Press (1982)

Erickson, T., Engel, J.: What goes before the CART? Introducing classification trees with Arbor and CODAP. Teach. Stat. **45**, S104–S113 (2023)

Garcia-Retamero, R., et al.: Using visual aids to improve communication of risks about health: a review. Sci. World J. 562637 (2012). https://doi.org/10.1100/2012/562637

Garcia-Retamero, R., Hoffrage, U.: Visual representation of statistical information improves diagnostic inferences in doctors and their patients. Soc Sci Med **83**, 27–33 (2013). https://doi.org/10.1016/j.socscimed.2013.01.034

Gigerenzer, G.: How to make cognitive illusions disappear: beyond "heuristics and biases." Eur. Rev. Soc. Psychol. **2**(1), 83–115 (1991)

Gigerenzer, G., Hoffrage, U.: How to improve Bayesian reasoning without instruction: frequency formats. Psychol. Rev. **102**, 684–704 (1995)

Gigerenzer, G., Hertwig, R., Hoffrage, U., Sedlmeier, P.: Cognitive illusions reconsidered. In: Handbook of Experimental Economics Results, vol. 1, pp. 1018–1034 (2008)

Hilbert, D.: Über das Unendliche. Mathematische Annalen **95**, 170 (1926)

Hoffrage, U., et al.: Communicating statistical information. Science **290**(5500), 2261–2262 (2000). https://doi.org/10.1126/science.290.5500.2261

Kurzenhäuser, S., Hoffrage, U.: Teaching Bayesian reasoning: an evaluation of a classroom tutorial for medical students. Med. Teach. **24**(5), 516–521 (2002)

Martignon, L., Hoffrage, U.: Wer wagt, gewinnt? Wie Sie die Risikokompetenz von Kindern und Jugendlichen fördern können. Hogrefe, Göttingen (2019)

Mousavi, S., Gigerenzer, G.: Risk, uncertainty, and heuristics. J. Bus. Res. **67**(8), 1671–1678 (2014)

Sedlmeier, P., Gigerenzer, G.: Teaching Bayesian reasoning in less than two hours. J. Exp. Psychol. Gen. **130**(3), 380 (2001)

Trevena, L.J., et al.: Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med. Inform. Decis. Mak. **13**(2), 7 (2013)

Venn, J.: On the employment of geometrical diagrams for the sensible representation of logical propositions. Proc. Camb. Philos. Soc. **4**, 47–54 (1880)

Wassner, C.: Förderung Bayesianischen Denkens — Kognitionspsychologische Grundlagen und didaktische Analysen. Franzbecker, Hildesheim (2004)