BOOSTING STUDENTS' SHARED KNOWLEDGE ON BASICS OF STATISTICS AND PROBABILITIES FOR UNDERSTANDING CRITICAL FACTS ABOUT COVID

Alexandra Kaiser and Laura Martignon
Institute of Mathematics Education, University of Education Freiburg, Germany
Department of Mathematics and Computer Science, Ludwigsburg University of Education, Germany
alexandra.sturm@ph-freiburg.de

The COVID-19 pandemic has taught us many lessons. One of them is that we need better tools for understanding the media and politicians. And we need tools for identifying wrong reasoning and wrong assertions on topics as urgent as tests, vaccines, and the associated risk reductions. We describe how we introduced instruments for better decision making on health issues, enhancing the traditional curriculum requirements of a standard course on stochastics for future teachers. This was just a pilot study. The first results are encouraging.

BACKGROUND

During the Middle Ages, people considered pandemics such as the Black Death, estimated to have killed at least one third of the population in Europe, as calamities sent from heavens to punish sinners. Centuries later during the Enlightenment, scientists discovered ways to measure calamities and estimate risks. Risks are namely detected by observing signs, symptoms, and characteristics, whose reliability can be assessed mathematically. With their discoveries, scientists paved the road towards an active, rational coping with such crises. In particular, medicine developed instruments and tests whose reliability could be estimated by means of probabilities, which had been introduced by mathematicians. In some people's views, calamities can be coped with by adopting rational procedures that make use of good probabilistic tools. However, as was discovered by cognitive scientists during the second half of the last century, coping with these mathematical instruments and making rational decisions is far from trivial when probabilities are presented formally (Gilovich et al., 2002; Tversky & Kahneman, 1974). We will describe the solution to these difficulties below.

Today, we face the COVID-19 pandemic in an epoch characterized by the tremendous effect of the media. With the advent of newspapers, television, and digital media, the spreading of measures and estimates of health risks has gone beyond doctors' offices and hospitals having an impact on public decision making. Many people are opting more and more to decide for themselves in situations involving public health risks, such as deciding whether to adopt a cure or vaccine. Everyone today follows television news and turns to the Internet. However, information in most media is abundant but sometimes wrong or incomplete and unclear. This is a new obstacle to rational decision making in times of COVID-19. One remedy is to boost people's risk literacy already in school, so that everyone becomes equipped with basic instruments for understanding health information in the media. Such literacy should be acquired and internalized during childhood and youth. Probability and statistics are now part of the school curricula basically across the whole world. There is a vast literature on the process that led to the introduction of these topics in school, but this is not the topic of our paper. We focus on the great progress in the research on representation formats and visualizations that facilitate the understanding and learning of probabilistic rules. These representation formats have been enriched by the advent of digital media and the possibility of adding dynamical components. The possibility of interacting with representations has enhanced the power of good visualizations.

The first group to investigate the representation formats that foster probabilistic reasoning developed around Gigerenzer, namely in his ABC Group during the nineties. Gigerenzer and Hoffrage (1995) tested participants' Bayesian reasoning based on natural frequencies instead of probabilities and reported a neat progress when compared to Kolmogorov probabilities. Mathematics educators reacted with enthusiasm and the development of adequate information formats for teaching probability and Bayesian reasoning has been impressive during the last quarter of a century. Our work is a direct consequence of these developments.

We report on a pilot study performed with a class of future teachers, in the promotion and implementation of educational tools for the interpretation and measurement of health risks in the context of the COVID-19 pandemic.

REPRESENTATION FORMATS FOR PROBABILISTIC SITUATIONS USED IN THE PILOT STUDY

In this section we briefly describe visualization formats used during the course for fostering probabilistic intuitions that appeared to be useful during the COVID-19 pandemic. For detailed treatments of visualizations and related strategies, see, for instance, Eichler, Böcherer-Linder, and Vogel (2020) and Gage and Spiegelhalter (2016). Formal probabilities and conditional probabilities were, of course, introduced, but, practically at the same time, translations into other formats were used during instruction. We list the main issues treated here, guided by the specific questions addressed.

Risk Reduction

During the COVID-19 pandemic, vaccine efficacy (VE) has often been communicated in terms of relative risk reduction instead of absolute risk reduction. Relative risk reductions express a percentage reduction in one group compared to another. Absolute risk reductions give the actual difference in proportions between one group and another. Thus, if a treatment reduces the incidence of a disease from "2 out of 100" to "1 out of 100," the relative risk reduction is 50%, which sounds huge. Observe Figure 1 on the right: The icon array on the left shows icons for 50 children who wear a helmet, whereas on the right side we see icons for 50 children who do not. Each head lying horizontally with a bandage on the face represents an injured child in a bike accident. Although the absolute reduction is from "30 out of 50" to "10 out of 50," the relative risk reduction is of 66.6 … %. The representation of risk reduction and risk enhancement by means of juxtaposed icon arrays, each representing one of the involved proportions, was introduced in Martignon and Hoffrage (2019).

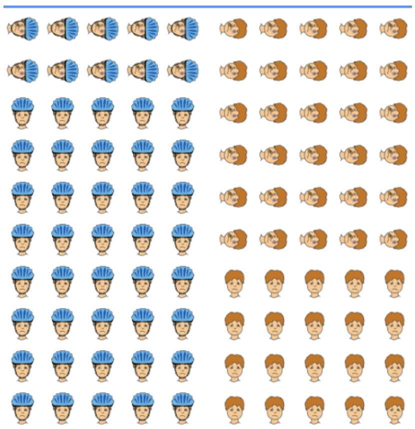

Figure 1. Risk reduction of injuries when using helmets: two icon arrays displaying corresponding proportions

Conditional Probabilities and Bayes' Formula

A formal way of treating probabilistic inference for updating the probability of a disease based on evidence is the following: assess the prior probability (also called base rate or incidence) that the disease (D) is present, and assess the probability that a certain piece of evidence (E) would be observed if the disease is present, or that it is not observed if the disease is absent. The corresponding formula is called Bayes' Rule, and it was first formulated by the mathematician, philosopher, and minister Thomas Bayes in the eighteenth century.

$$P(D|E) = \frac{P(E|D) \cdot P(D)}{P(E|D) \cdot P(D) + P(E|D^C) \cdot P(D^C)}$$

The formula shows how to solve the important evidential reasoning problem: assess P(D|E), that is, the probability that the patient has the disease given that they tested positive on the test. However, people are notoriously bad at manipulating probabilities, as a plethora of empirical studies have shown. The first seminal study by the medical doctor Eddy was published in the collection of papers on heuristics and biases published by Gilovich, Tversky, and Kahneman in 1982 (Eddy, 1982; Gigerenzer & Hoffrage, 1995). In Eddy's classical study on doctors' estimate of the probability that a certain disease is present given that a test of the disease is positive, he discovered that his participants made mistakes based on misconceptions. The so-called predictive value of the test was estimated as being close to the chances of the test detecting the disease. Additional tools that we introduced to help students reason about problems such as the one presented by Eddy (1982) include 2 x 2 contingency tables, unit squares, icon arrays, natural frequencies, and double trees.

2 x 2 Contingency Tables

A 2 x 2 table is a traditional and well-known visualization tool that summarizes data. Frequency data are cross classified according to two categorical variables.

Unit Squares

The area of the unit square visualization corresponds to the size of a sample, or 100%. The first step for using the unit square to estimate the probability that a certain disease is present given that a test of the disease is positive is to subdivide the square according to the base rate or incidence of the disease. See the vertical division for the unit square displayed in Figure 2. The second step is to subdivide each rectangle (left and right) horizontally according to sensitivity/specificity. See the horizontal segments added dividing sections of the rectangles in the unit square in Figure 2. Thus, the square is subdivided into four rectangles representing the four intersections, namely: infected and positive test, not infected and positive test, infected and negative test, and not infected and negative test. It is interesting to let the students play with a plugin such as one available for GeoGebra. Students could drag and vary the base rate and the corresponding rectangle (https://www.geogebra.org/m/uzpmt8gj). A slight disadvantage of the unit square occurs when the base rate is very small because the first rectangle is almost invisible.

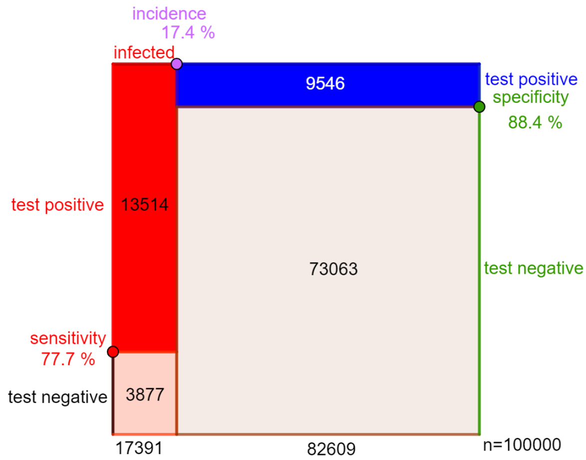

Figure 2. Unit square subdivided in rectangles according to base rate, sensitivity, and specificity

Icon Arrays, Natural Frequencies, and Double Trees

Our next step with students was to introduce natural frequencies and icon arrays for concrete examples of diseases and tests. Natural frequencies translate probabilities into proportions; thus, a probability of 0.09 can be expressed as "9 out of 100." These two expressions are not mathematically equivalent. Natural frequencies can be seen as "expected frequencies" in the sense that they translate a probability of, say, 0.09 into the frequency we would expect in 100 randomly selected cases (see, for instance, Gage & Spiegelhalter, 2016). An icon array is an organized display of icons representing individuals of a population such as the visual presented in Figure 1 to consider the risk reduction associated with wearing a bicycle helmet. Icons are distinguished by color, shading, shape, or form and draw on people's natural disposition to count and allow for identification.

Exercises Treated by Students and Discussed During the Exercise Session

The following list summarizes the concrete examples introduced to students in exercises.

- Description of a diabetes test (from a school textbook);
- Antibody test against SARS-Cov-2 from the Böcherer-Linder and Sturm (2021) article on the influence of incidence on positive predictive value and negative predictive value;

- The probability of myocarditis after vaccination;
- Negative tests for high incidences (e.g., the city in which the university is located, Freiburg, Germany);
- Predictive power of a positive COVID-19 test in Freiburg;
- Technical medical terms explained;
- "43% of all new infections that are symptomatic are now among vaccinated individuals" (Sahra Wagenknecht of the Left Party in a TV Talk-show, https://www.ardmediathek.de/video/anne-will/steigende-neuinfektionen-sorge-wegen-impfskepsis-hilft-oder-schadet-mehr-druck-auf-ungeimpfte/das-erste/Y3JpZDovL25kci5kZS80ZDlmMDU2ZS1kNmI0LTQwYmQtOWMwYy02M2FkYjU4NDA4OGU); and
- A newspaper article on "Sloppiness in rapid tests for schoolchildren." (https://www.lkz.de/lokales/stadt-ludwigsburg_artikel,-bis-zu-70-prozent-der-positiv-getesteten-schueler-in-ludwigsburg-sind-doch-negativ-_arid,630296.html)

For the illustration of Bayesian tasks, we used different visualizations and representation formats that proved to be helpful for understanding and determining solutions in the sense that they enhanced students' performance: tree diagrams (simple and double trees), unit squares, 2x2-tables, as well as the systematic use of natural frequencies and icon arrays. See Figure 3.

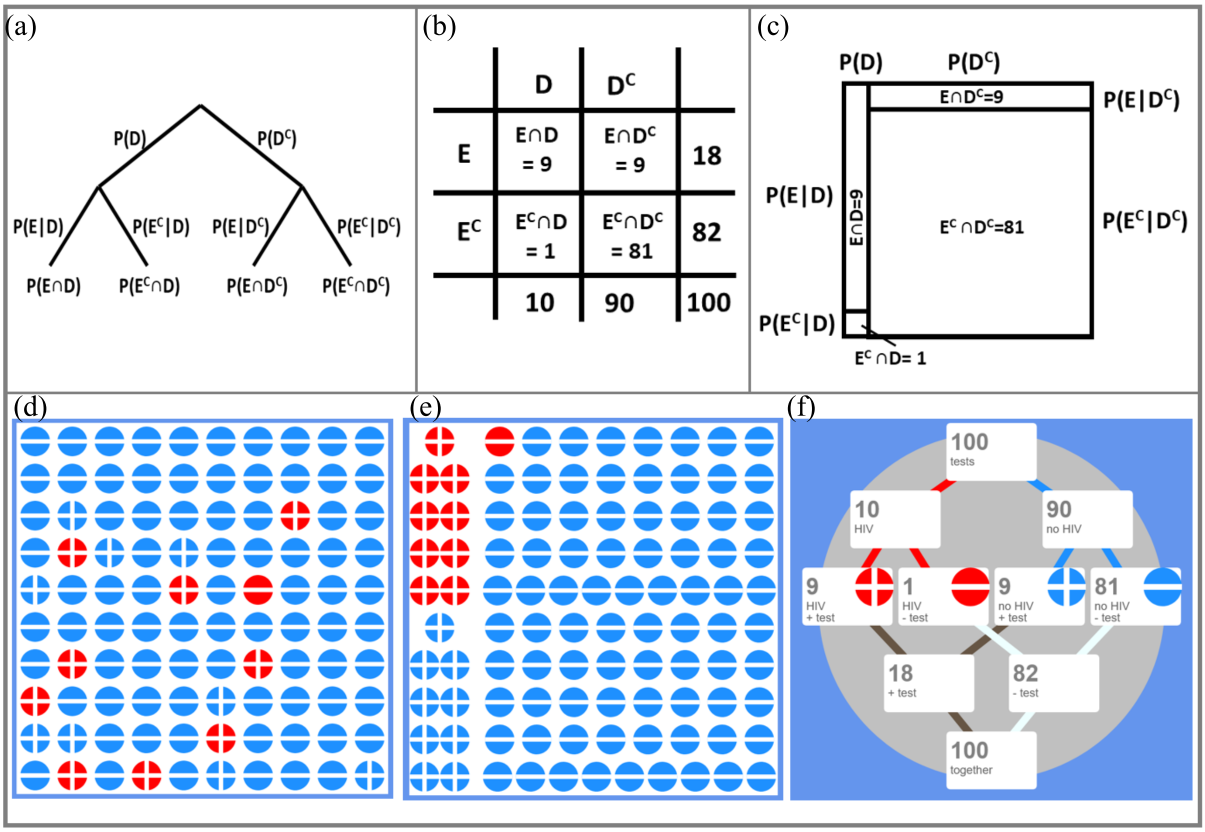

Figure 3. (a) Typical probabilistic tree for Bayesian reasoning, (b) 2x2 table, (c) unit square, (d) unsorted icon array, (e) sorted icon array, and (f) double tree corresponding to those data

OUR PROCEDURE IN A PILOT STUDY

Our work was developed during a course on basic probability and statistics (Daten und Zufall) with future teachers at the University of Education in Freiburg during the Winter-term 2021–2022. The instruction units required four hours each week, and the practical sessions (exercises units) required two hours per week. Topics and questions for each practical session were discussed and chosen by the two authors based on the topics covered during the instruction hours.

We made a conscious search for examples from media news or for information on COVID-19 related issues. For example, a highly visited semi-official source reporting on COVID-19 diagnostic tests stated: "If a test ends with the result 'positive,' the tested person is infected with the virus, 'negative' means that he is not infected." (https://www.ndr.de/ratgeber/gesundheit/Corona-Selbsttest-und-Schnelltest-So-funktionieren-sie,schnelltest226.html). But a positive test does not imply necessarily that the person has contracted the virus.

Our main action was to describe facts such as these from the news, be it from newspapers, television, or websites, as examples that called for treatment with the methods treated in the course. Notably, we also worked with citations from statements of politicians. In the discussions about the vaccine against COVID-19, even politicians insisted that vaccines are superfluous if, after all, more than 45% of Intensive Care Units at hospitals are assigned to vaccine patients.

We also designed exercises with tasks that made use of the learned tools for answering specific questions related to the pandemic. We basically used content that was very present in the media in Germany or locally in the city of Freiburg.

We list some characteristics of our activities.

- Students were instructed to work with dynamic plugins.
- Students explored risk using the dynamic webpage, "Worth the risk" (http://www.eeps.com/projects/wwg/wwg-en.html).
- Students engaged with tasks, and the typical task asked students to interpret a visualization and calculate the probability that someone is healthy, given that they test positive.

Exam Topics

Questions asked on examinations covered a wide spectrum of topics, which is typical for a traditional stochastics course in that level. Two tasks were devoted to the issues on COVID-19 treated during the course.

RESULTS

The results we describe here are those from the final examination after students completed the course. Of the 54 students in the class, 48 participated in the research. Among the questions on typical topics in stochastics, we included two specific questions tightly connected with COVID-19. See Figure 4. Another task was to describe typical errors in the media concerning facts about COVID-19.

> Assume that the sensitivity and specificity of a test remain constant at 0.96 while the incidence of the disease is enhanced to its double, triple and quadruple
> (a) How does the positive predictive power change? Does it increase or decrease?
> (b) How does the negative predictive power change? Does it increase or decrease?
>
> Now assume the incidence is reduced to 1/100 of its original value.
> (c) How does the positive predictive power change? Does it increase or decrease?
> (d) How does the negative predictive power change? Does it increase or decrease?

Figure 4. Final examination task administered on February 17, 2022

We were positively impressed by students' results, which we had not expected. Two graphs of students' results are presented in Figure 5. For Figure 5(a), the first bar indicates the number of correct students' answers on the task displayed in Figure 4 regarding the impact of incidence changes on positive and negative predictive values. The second bar presents results for the question on examples of severe mistakes in the media. Figure 5(b) illustrates students' results for both tasks combined. One aspect seemed relevant: students had used tree diagrams more often than other visualizations during the exercise sessions. However, for the task on varying incidences in the context of COVID-19 tests in the final exam, most students used the unit square representation. In fact, during the exercise sessions, students had mentioned their ease with this particular representation. Students' statements and their performance seemed to corroborate our impression that their work with the interactive plugin during the exercise session fostered their intuitions and enhanced their grasp of the varying predictive values in the task.

While closing this section we would like to report that, in general, students were interested in acquiring tools for reasoning and making decisions about COVID-19. Their cooperation was impressive. We acknowledge, in particular, useful and helpful comments by Ch. Hermann, N. Triebs, G. Wittmann and S. Winkler.

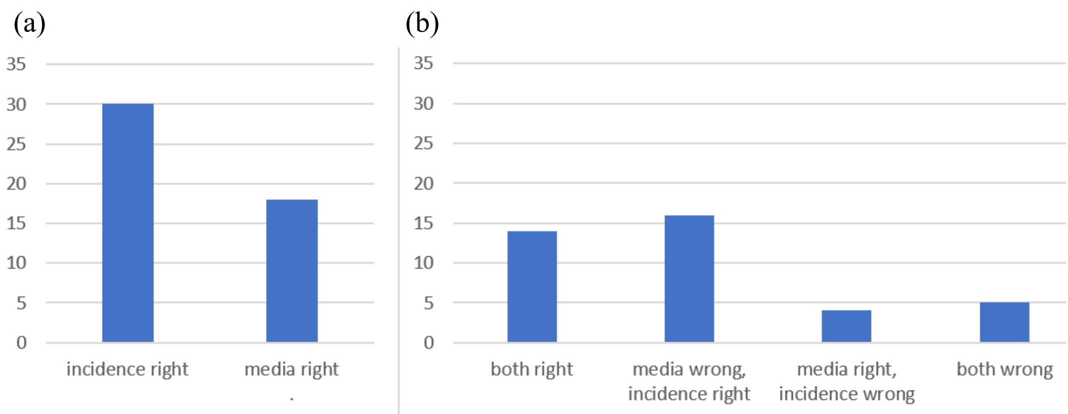

Figure 5. (a) Results for the Figure 4 task and the media task and (b) results for the two tasks combined

CONCLUSION

The COVID-19 pandemic has been a difficult and painful period for everyone. Besides all the events involving suffering and despair, the pandemic has also shown how some basic tools from probability and statistics can become handy. In fact, these tools can foster understanding of information about tests, vaccines, and treatments in general. These tools can be used when deciding whether to undergo a certain treatment or whether to get a vaccine and thus can be essential for saving lives. They should be part of the training of future teachers and become standard components of the school curriculum. In our pilot study, we gained the conviction that this is possible, and that it fosters future teachers' enthusiasm for stochastics.

REFERENCES

Böcherer-Linder, K., & Sturm, A. (2021). Mit einheitsquadraten statistisches verständnis fördern [Encourage statistical understanding with unit squares]. *Mathematik Lehren, 224*, 28–31.

Eddy, D. (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 University Press. https://doi.org/10.1017/CBO97805118809477.019

Eichler, A., Böcherer-Linder, K., & Vogel, M. (2020). Different visualizations cause different strategies when dealing with Bayesian situations. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.01897

Gage, J., & Spiegelhalter, D. (2016). *Teaching probability*. Cambridge University Press.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704.

Gilovich, T., Griffin D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Martignon, L., & Hoffrage, U. (2019). *Wer wagt, gewinnt? Wie sie die risikokompetenz von kindern und jugendlichen fördern können* [Who dares, wins? How you can promote the risk competence of children and young people]. Hogrefe.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124