13th International Congress on Mathematical Education
Hamburg, 24-31 July 2016

THE IMPACT OF VISUALIZATION ON UNDERSTANDING CONDITIONAL PROBABILITIES

Katharina Böcherer-Linder, Andreas Eichler, Markus Vogel

University of Education Freiburg, University Kassel, University of Education of Heidelberg

The aim of our research is to compare the efficiency of two visualizations in the field of conditional probabilities, i.e. the tree diagram and the unit square with natural frequencies. In this paper we report the method of our study including a test instrument that we administered to 148 prospective teachers. We further discuss results of our study that give evidence that the unit square is more efficient gain both procedural knowledge and conceptual knowledge referring to the Bayes rule.

INTRODUCTION

Conditional probabilities are a subject in which the human intuition often seems to be misleading (e.g. Kahneman & Tversky, 1972). However, research gained strong evidence that the way of presenting the needed information plays a crucial role (c.f. Gigerenzer & Hoffrage, 1995). Moreover, the way of visualizing conditional probability has a considerable impact on students learning. Particularly two strategies were identified to facilitate students' learning, i.e. the tree diagram with natural frequencies and the unit square (Sedlmeier & Gigerenzer 2001, Bea 1995, Wassner 2004).

Sedlmeier and Gigerenzer (2001) investigated the ability of adults (students at university) to solve tasks using the Bayes' rule. They found that the tree diagram with natural frequencies (fig. 1) is a more efficient visualization than a common tree diagram with probabilities. Sedlmeier and Gigerenzer (2001) gave two reasons explaining the efficiency of the tree diagram with natural frequencies. The first reason concerns the way of information representation based on natural frequencies that facilitates Bayesian reasoning (c.f. Gigerenzer & Hoffrage, 1995). The second reason refers to the assumption that the tree represents a powerful tool to structure given information because of his sequential and hierarchical structure. In the research of Sedlmeier and Gigerenzer (2001), the efficiency of students' learning was measured by the number of students' correct solutions of tasks for which applying the Bayes' rule is necessary. Thus, this study primarily addressed students' procedural knowledge (Hiebert & Carpenter, 1992). In addition, Binder et al. (2015) replicate the finding that natural frequencies are more efficient to solve problems referring to the Bayes rule. They further find that there is no difference in the efficiency of using a tree with natural frequencies or a 2 x 2-table as visualization of the statistical information.

Further, Bea (1995) yielded evidence that using a unit square is more efficient than using the tree diagram with probabilities when students' learning of conditional probability was regarded. In his research, Bea used the unit square with probabilities. The unit square visualizes conditional probabilities and Bayes' rule geometrically based on proportions of line segments and proportions of partial areas of the square. Accordingly, Bea (1995) showed that the unit square particularly had promoted students' conceptual knowledge (Hiebert & Carpenter, 1992).

However, the aforementioned research results do neither involve a direct comparison referring to the impact of these both visualizations on students' learning nor a direct comparison by only using natural frequencies (fig. 1). Based on the considerations above, our research focuses on the comparison of these two strategies of visualization using throughout natural frequencies and on the definition of that part of students' knowledge that is promoted by certain visualization.

VISUALIZING CONDITIONAL PROBABILITIES

We explain the way how we visualize conditional probabilities by showing an example:

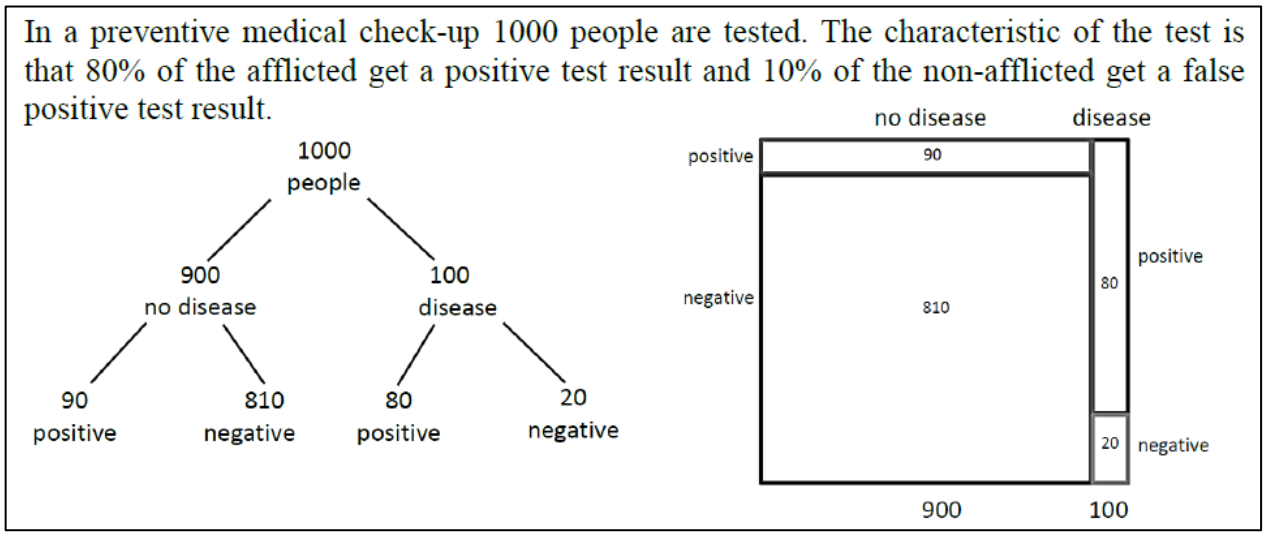

Fig. 1: The tree diagram and the unit square with natural frequencies.

By using natural frequencies, the conditional probability $p(disease \mid positive)$ can be obtained by calculating the ratio of the number of people being afflicted and tested positive (80) divided by the number of all people tested positive (90+80).

In Germany, only the tree diagram with natural frequencies plays a substantial role in statistics education, the unit square is almost unknown. However, we could show in our former research, that the unit square with natural frequencies is equally efficient to extract relevant information of statistical situation (Böcherer-Linder et al., 2015). This was an important result, because it is now possible to compare the results of other tasks that are more complex.

METHOD AND HYPOTHESES

The aim of our study was to compare the efficiency of the two visualizations to promote primarily procedural knowledge or primarily conceptual knowledge in the field of conditional probabilities. We designed test items that address procedural knowledge by showing the tree diagram or the unit square respectively and asking to calculate conditional probabilities in form of ratios, e.g. "Calculate the percentage of people afflicted with the disease among the people tested positive". To design test items that address conceptual knowledge we showed a diagram with the relevant information and used tasks concerning the influence of the base rate (for an example see fig. 2). All in all, we had three test items that were structurally identical. Four sub-items of each item addressed conceptual knowledge and one sub-item addressed procedural knowledge. We had two different

versions of the test. One showed tree diagrams to present the needed information, the other showed unit squares. The tasks were identical, only the visualization differed.

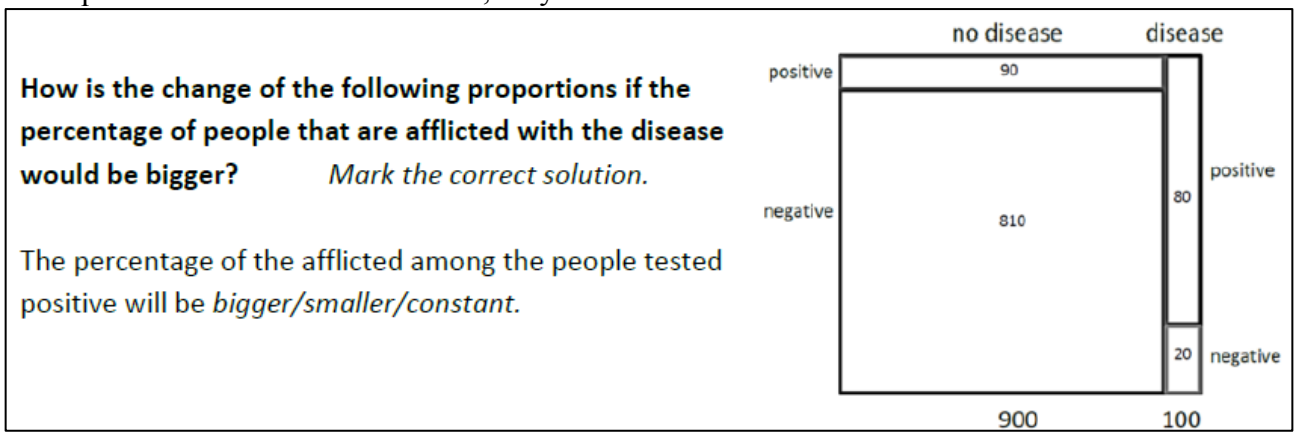

Fig. 2: Example of an item that address conceptual knowledge showing the unit square.

Concerning the test items addressing procedural knowledge our first hypothesis was: The tree diagram with natural frequencies is more efficient because of his hierarchical and sequential structure (Wassner, 2004). Concerning the test items addressing conceptual knowledge our second hypothesis was: The unit squar with natural frequencies is more efficient for understanding the influence of the base rate, because - beyond the same numerical information given within both kinds of visualization - the proportions are (at least partially) displayed within in the unit square but not within the tree diagram (Eichler & Vogel 2010).

The study took place at the University of Education Heidelberg in 2015. The sample consists of 148 prospective teachers that are not in the first year of their study (for the reason of this restriction see Böcherer-Linder et al., 2015). The sample was randomly divided into two subsamples. The participants got a brief description of a visualization based on a simple example, the first subsample referring to the use of the tree diagram with natural frequencies, the second sample referring to the unit square with natural frequencies. Both descriptions were identical as regards stochastic content, they differ only as regards the kind of visualization.

RESULTS

Fig. 3 shows the average of correct answers for the three sub-items concerning procedural knowledge. In contrast to our first hypothesis, the unit square was more efficient. The difference

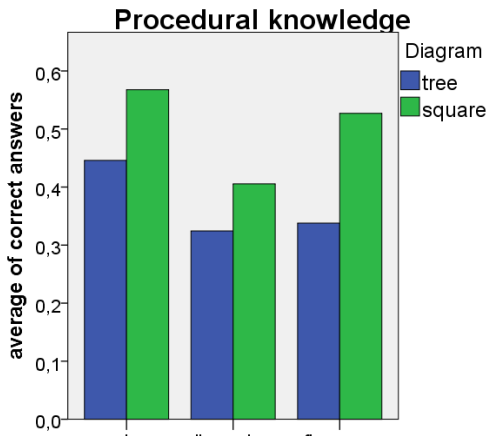

Fig. 3

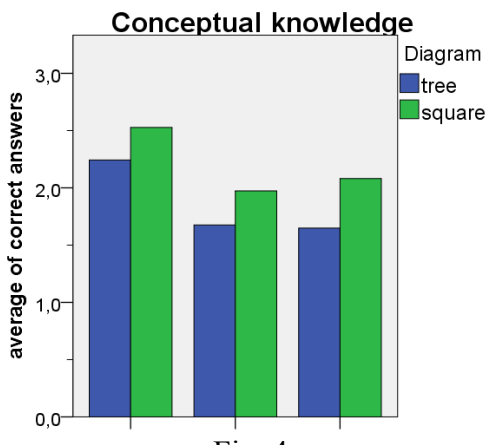

Fig. 4

between the mean-values summarized over the three sub-items is significant (p=0,044).

In fig. 4 we provide the results concerning conceptual knowledge. Each of the three test items had four sub-items addressing conceptual knowledge. We summarized over the sub-items and show in fig. 4 the average of correct answers. Our second hypothesis was confirmed: regarding the mean-value of correct answers summarized over the sub-items the p-value is equal to 0,042. Hence, for both kinds of knowledge, we have statistical evidence that the unit square is more efficient.

DISCUSSION

Our results can have great implication on the communication of risk and on the teaching and learning of conditional probabilities. Actually, the tree diagram with natural frequencies is often promoted as an outstanding way to visualize conditional probabilities (e.g. Spiegelhalter & Gage, 2014). However, our research gives evidence that the unit square with natural frequencies is more efficient for the calculation of conditional probabilities in form of ratios and for the understanding of the influence of the base rate.

References

Binder, K., Krauss, S. & Bruckmaier, G. (2015). Effects of visualsing statistical information – an empirical study on tree diagrams and 2 x 2 tables. *Frontiers in psychology*, 6:1186. doi: 10.3389/fpsyg.2015.01186.

Böcherer-Linder, K., Eichler, A. & Vogel, M. (2015). Understanding conditional probability through visualization. In H. Oliveira, A. Henriques, A. P. Canavarro, C. Monteiro, C. Carvalho, J. P. Ponte, R. T. Ferreira & S. Colaço (Eds.), *Proceedings of the International Conference Turning data into knowledge: New opportunities for statistics education*. Lisbon, Portugal: Instituto de Educação da Universidade de Lisboa.

Bea, W. (1995). *Stochastisches Denken* (Stochastical reasoning). Frankfurt a.M.: Lang.

Eichler, A., & Vogel, M. (2010). Die (Bild-)Formel von Bayes (The (picture) formula of Bayes). *PM - Praxis der Mathematik in der Schule, 52(32),* 25-30.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency Formats. *Psychological Review, 102 (4)*, 684-704.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: Macmillan.

Kahnemann, D., & Tversky, A. (1972). Subjective Probability: A judgement of representativeness. *Cognitive Psychology, 3(3),* 430-454.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130(3)*, 380-400.

Spiegelhalter, D., & Gage, J. (2014). What can we learn from real-world communication of risk and uncertainty? In K. Makar (Ed.) *Proceedings of the Ninth International Conference on Teaching Statistics*. Flagstaff, USA: International Statistical Institute an International Association for Statistical Education. Retrieved from http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_PL2_SPIEGELHALTER.pdf.

Wassner, C. (2004). *Förderung Bayesianischen Denkens. Kognitionspsychologische Grundlagen und didaktische Analysen* (Promoting Bayesean thinking. Basics referring to cognitive psychology and educational analysis). Hildesheim: Franzbecker.