FREQUENCY PHOBIA IN SPITE OF PROBABILITY BLINDNESS

Patrick Weber, Stefan Krauss, & Karin Binder
Mathematics Education, University of Regensburg, Germany
Patrick.Weber@ur.de

Empirical research has repeatedly proven the facilitating effect of "natural frequencies" when it comes to solving Bayesian problems (e.g., Gigerenzer & Hoffrage, 1995). However, in a study with 183 university students, we found that the beneficial strategy of translating complicated conditional probabilities (e.g., "80%") into natural frequencies (e.g., "8 out of 10") is for the most part not being used by participants. On the contrary, even when statistical information was presented in natural frequencies, most of the students translated the natural frequencies they had been given back into probabilities – with the unfortunate consequence that they could no longer solve the task.

BACKGROUND

For more than 20 years, many empirical studies, starting with a seminal article by Gigerenzer & Hoffrage (1995), have contributed to the now widespread acceptance of natural frequencies (e.g., "8 out of 10") as a means of facilitating Bayesian reasoning tasks. Compared to conditional probabilities (e.g., "80%"), tasks presented in this information format yield much higher performance rates (for an example of a Bayesian reasoning task, see Table 1; left: probability format; right: natural frequency format). A recent and comprehensive meta-analysis (McDowell & Jacobs, 2017) shows that, whereas only 4% of participants were able to solve a Bayesian reasoning task formulated with probabilities, performance rates on problems presented in natural frequency format were significantly higher (24%). However, even this increase in performance seems restricted, bearing in mind the extensively investigated beneficial effects of intuitive natural frequencies. Why, then, can *only* 24% of participants on average draw the correct inferences? While previous research focused mainly on explaining, discussing, and replicating the positive effects of natural frequencies (e.g., Binder et al., 2015; Siegrist & Keller, 2011), we tried to shed light on why over three quarters of participants still failed to solve the Bayesian reasoning task when it was given in natural frequency format.

Therefore, we are interested in understanding what participants who fail in natural frequency versions actually do when working on a Bayesian problem. In order to approach this issue, we first conducted a small schoolbook review of a sample of typical Bavarian math texts and workbooks. The dominating format in sections on statistics is, of course, probabilities. However, we found that even many of the tasks presented in natural frequency format were not actually solved by the textbooks with natural frequencies but instead with probabilities (e.g., Freytag et al., 2008). This seems rather surprising, since especially in Germany, there have been many recent efforts to establish the natural frequency concept not only in schools but also in various other fields, for example, in medicine (Gigerenzer et al., 2007). Similarly, universities offer many supplemental courses for high-school mathematics teachers in which opportunities for teaching natural frequencies are presented and discussed. Moreover, for the very first time, the new Bavarian curriculum (ISB, 2016) contains a section on teaching natural frequencies. Nevertheless, teachers and students alike seem to be more familiar with probabilities than with natural frequencies.

Many psychological theories have addressed the effects of strong familiarity with a particular concept or strategy on problem-solving. While such a preestablished strategy facilitates finding a solution to problems of a specific type, it hinders insight in situations where a different strategy might be more efficient. This phenomenon of a fixed mental set "blocking" relevant solution strategies is called the *Einstellung* effect and was first detected by Luchins in his famous water-jar experiments (1942). While in these experiments, participants learned a fixed solution strategy directly before they had to solve a task with a different strategy (which most of them were unable to do), the Einstellung effect was recently expanded to situations where strategies that had been learned over time "blinded" participants, so to speak, to a simpler solution. For instance, experienced chess players were unable to see a simpler way to checkmate their opponents because

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

this simpler but unusual solution was blocked by a common but more moves-demanding way of achieving checkmate (Bilalić et al., 2008).

Table 1. Problem formulations for the heroin addiction problem

	Heroin addiction problem	
	Probability version	Natural frequency version
Intro-duction	Imagine that you randomly meet a person with fresh needle pricks in the street. You are interested in whether this person is addicted to heroin.	
Version	On the internet, you find the following information for a sample of 100,000 people: The probability that one of these people is addicted to heroin is 0.01 %. If one of these people is addicted to heroin, the probability is 100 % that he or she will have fresh needle pricks. If one of these people is not addicted to heroin, the probability is 0.19 % that he or she will nevertheless have fresh needle pricks.	On the internet, you find the following information: 10 out of 100,000 people are addicted to heroin. 10 out of 10 people who are addicted to heroin will have fresh needle pricks. 190 out of 99,990 people who are not addicted to heroin will nevertheless have fresh needle pricks.
Visual aid	• actively constructed tree diagram or • tree diagram (prob.)	• actively constructed tree diagram or • tree diagram (nat. freq.)
Question	What is the probability that one of these people is addicted to heroin, if he or she has fresh needle pricks?	Of the people who have fresh needle pricks, what is the proportion of them addicted to heroin?

Taken together, the schoolbook review and the considerations on the Einstellung effect gave us a first hint regarding why so many participants still fail in Bayesian reasoning problems when natural frequencies are presented: lacking familiarity with the natural frequency concept might provoke participants to switch back to preestablished but unintuitive probabilities (i.e., their *calculation format* differs from the *presentation format*). Therefore, we had a closer look at questionnaires used in previous studies on Bayesian reasoning (conducted by Binder et al., 2015, and Krauss et al., 1999). We thus confirmed our impression that many participants did not actually use natural frequencies for their calculations but instead translated them back into the less intuitive probability format, which might explain why three-quarters of participants still failed with natural frequencies (McDowell & Jacobs, 2017).

If this is indeed the explanation for that restricted performance in Bayesian reasoning tasks in natural frequency format, what can be done to prevent participants from giving in to their *Einstellung* and translating intuitive frequencies into complicated probabilities? One idea is to encourage participants to more thoroughly process the information in natural frequency format in order to increase the chance that they actively use a natural frequency algorithm for their solution. This could be achieved by making the given natural frequencies more salient, for example by providing not only a textual version of the task but a corresponding visualization that also includes natural frequencies. With the verbal information in the task supported by visual information, the two different cognitive-processing channels (verbal and visual) would be activated, thus enabling a deeper understanding of the problem (Paivio, 1990; Mayer, 2005). On the other hand, participants' engaging with the task in frequency format could even be enhanced by having them construct a visualization *on their own* (Cosmides & Tooby, 1996; Micallef et al., 2012) so that they *actively* deal with the visual information. Both situations (*presenting* a visualization and actively *constructing* one) might encourage participants to actually stay with natural frequencies instead of relying on the familiar probabilities.

In sum, we wanted to examine systematically if participants prefer calculating with probabilities over natural frequencies, and what impact this potential preference has on performance rates. Moreover, we were interested in whether either a presentation or an active construction of an additional visualization based on natural frequencies would prevent participants from falling back into well-known but unintuitive probabilities.

EXPERIMENTAL STUDY

We conducted an empirical study with N=114 students from the University of Regensburg (Bavaria) in 2016. After having detected significant effects regarding the preference of calculating with probabilities over natural frequencies, we decided to extend our sample by another N=69 students in order to strengthen the results obtained. All participants were enrolled in different study programs, with the majority studying some kind of mathematics education.

We chose tree diagrams as visualizations in our study. Participants had all completed German high school (*Gymnasium* or *FOS/BOS*), so they were all familiar with tree diagrams as effective visualizations for different problems in probability theory (see, e.g., the Bavarian curriculum: ISB, 2016). Moreover, tree diagrams can be equipped with both natural frequencies and probabilities (see Figure 1), which enabled us to control for possible interactions of presentation format and visualization.

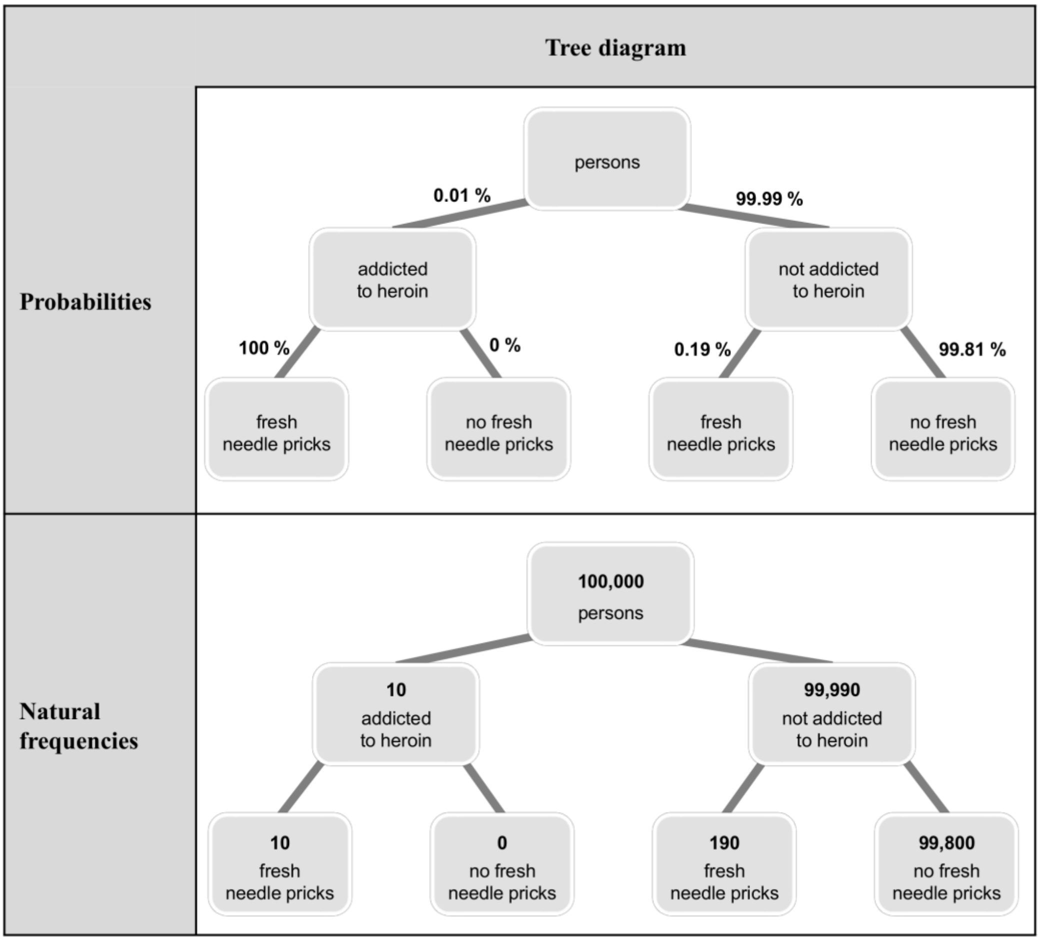

Figure 1. Probability tree and natural frequency tree for the heroin addiction problem

Since we wanted to test two presentation formats (natural frequencies vs. probabilities), two problem contexts (heroin addiction problem vs. car accident problem), and two ways of engaging with a visualization (using a presented tree diagram vs. actively constructing one), we implemented a 2×2×2 design. Thus, each participant received two tasks, one in probability and the other in natural frequency format, with one of them concerning the heroin addiction problem and the other the car accident problem. We chose these two problem contexts because they are not as well known as, for example, the famous mammography problem, so that there was very little chance of a participant already having heard of the problems. Table 1 shows problem formulations for the heroin addiction problem both in probability and natural frequency presentation formats. We systematically varied the contexts as well as the presentation formats. In one of the two tasks, participants were asked to construct a tree diagram on their own, whereas a fully drawn tree diagram was presented along with the problem formulation in the other task.

In order to judge precisely whether a participant had mainly applied probabilities or natural frequencies for his or her calculations, we explicitly asked participants to write down every step they took to obtain their estimate. Moreover, we triggered the active construction of the tree diagram by an extra prompt (“Please draw a tree diagram illustrating the given problem.”) after the problem formulation.

RESULTS

Generally, the beneficial effect of *presenting* a task in natural frequencies was replicated. Tasks that were presented in probability format had performance rates of 21%, whereas tasks formulated with natural frequencies were solved correctly in 36% of the cases. These performance rates seem rather high compared to other studies (McDowell & Jacobs, 2017) and could be explained by our sample that mainly consisted of students of mathematics education. A new finding is that participants showed a general preference for probabilities over natural frequencies. If a task was presented in probability format, 78% made use of probabilities for their calculations, and even if the presentation format was in natural frequencies, almost half of the participants (48%) nevertheless chose to translate the frequencies back into probabilities (see Figure 2). Overall, almost two thirds (65%) of participants applied probabilities when working on a given task.

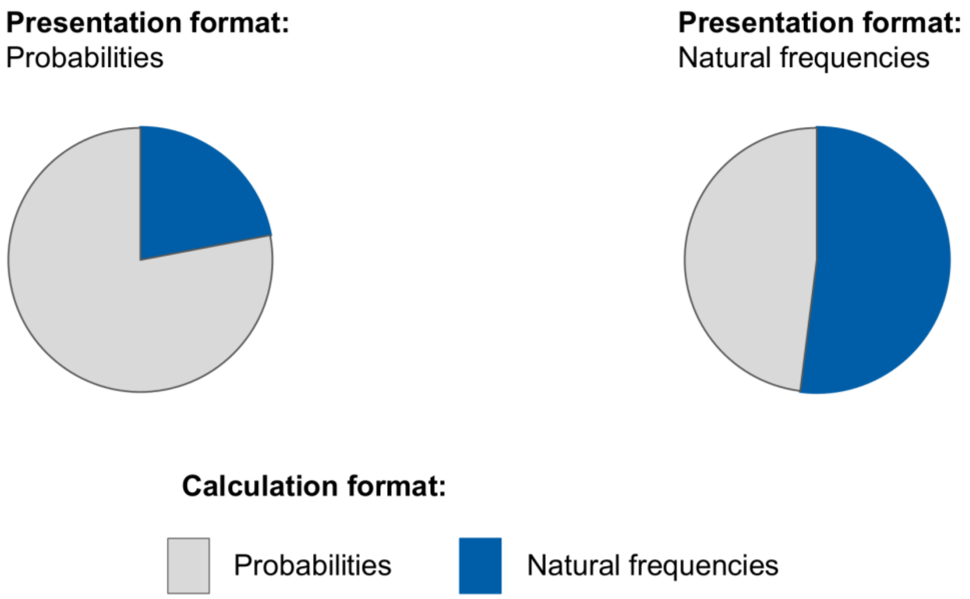

Figure 2. Proportions of probability algorithms and natural frequencies algorithms used by participants depending on the presentation format (left: probabilities; right: natural frequencies)

This preference for unintuitive probabilities over natural frequencies had significant consequences on the participants’ performance: When they applied natural frequencies, performance rates were as high as 58%, but if the calculation format was in probabilities, performance rates shrank to only 12% averaged over both problem contexts. Figure 3 shows the percentages of correct Bayesian inferences for all four problem versions, depending on whether the calculation format was in probabilities or in natural frequencies. These results indicate that the *calculation format* has not only a significant effect on performance rates, but that it is an even stronger predictor of performance in Bayesian reasoning tasks than the *presentation format* is.

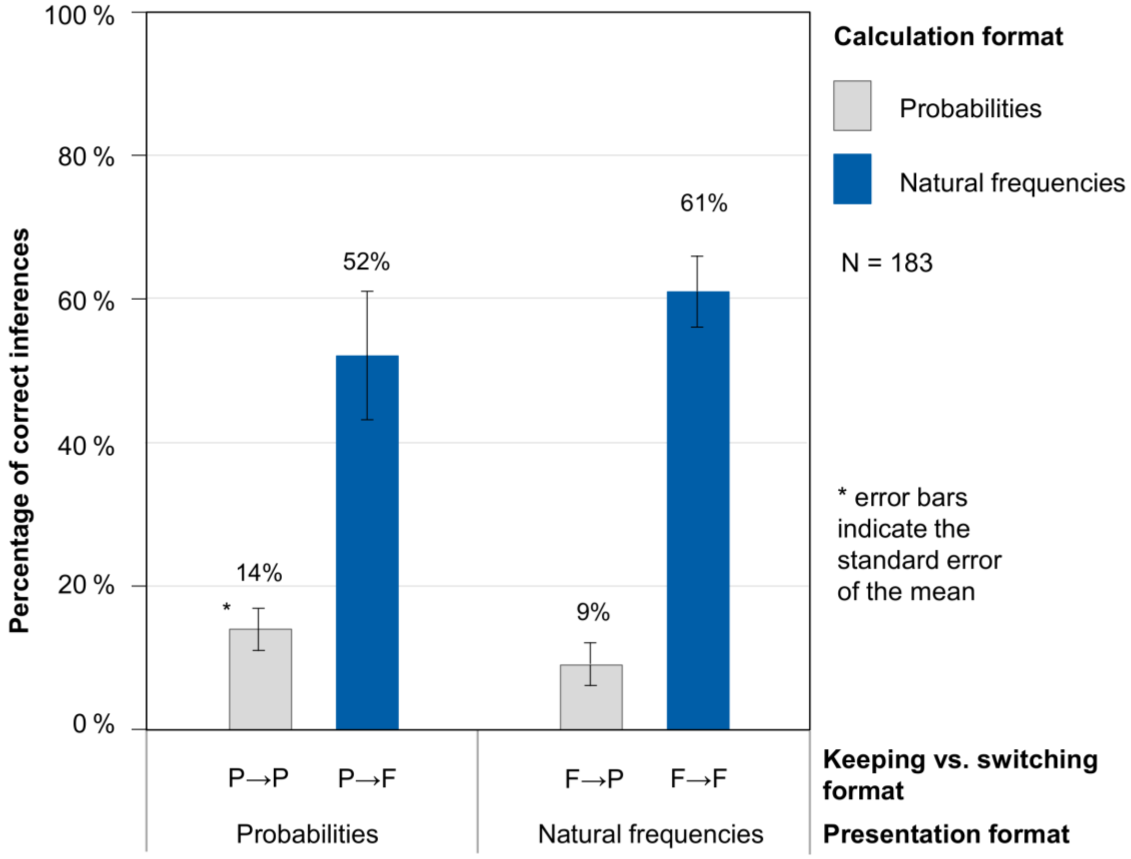

Figure 3. Percentages of correct inferences in all four problem versions (heroin addiction problem vs. car accident problem presented in probability vs. frequency format)

Our results did not show any significant indications as to whether the presentation or the active construction of a tree diagram is more likely to prevent participants from falling back into the probability format. Because of our study design, we do not know how many participants would have translated back into probabilities if they had not engaged with a visualization at all. However, note that our design is very conservative because we assume that even more participants would have chosen to calculate with probabilities if no visualization had been given or constructed.

CONCLUSION

In an empirical study on Bayesian reasoning with N=183 students from the University of Regensburg, we found that the majority of participants do not actively use natural frequencies in Bayesian reasoning tasks. Even if the task is presented in the intuitive natural frequency format, about half of the participants still prefer calculating with probabilities instead. This has a negative impact on performance rates: participants working with probabilities perform significantly worse than those who apply frequencies for their calculations. Moreover, the *calculation format* is an even stronger predictor for performance than the *presentation format*, on which previous research had mainly concentrated (e.g., Barbey & Sloman, 2007; Siegrist & Keller, 2011).

Our findings imply that the concept of natural frequencies is not yet as well established in schools and universities as would be necessary to make students actually use frequencies in Bayesian reasoning situations. Even though research has shown that for teaching statistics, the frequency concept can be a powerful tool for fostering insight into Bayesian reasoning problems, recent efforts to implement natural frequencies in high school and university curricula do not appear to be enough, since students seem to avoid natural frequencies as if they have a phobia about them. Therefore, we suggest a stronger implementation of natural frequencies in high school and university statistics education as well as in teacher training.

The extent to which an implementation of natural frequencies in statistics education is reasonable remains an open question, since their use is limited to specific types of problems (e.g., Bayesian reasoning or cumulative risk judgment; see McCloy et al., 2007). Future research on this topic might further investigate to what extent current teachers are familiar with the frequency concept in order to determine if the natural frequency format needs to be a stronger focus in teacher training. Moreover, it would be interesting to determine effective methods (other than presenting or constructing visualizations) that would prevent people from falling back into probabilities in Bayesian reasoning tasks.

REFERENCES

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: from ecological rationality to dual processes. *Behavioral and Brain Sciences, 30,* 241–297.

Bilalić, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, *108*(3), 652–661.

Binder, K., Krauss, S., & Bruckmaier, G. (2015). Effects of visualizing statistical information: An empirical study on tree diagrams and 2 × 2 tables. *Frontiers in Psychology*, *6*(1186).

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*(1), 1–73.

Freytag, C., Herz, A., Kammermeyer, F., Kurz, K., Peteranderl, M., Schmähling, R., Schmitt, B., Sinzinger, M., Zebhauser, E., & Zebhauser, M. (2008). *Fokus Mathematik 10 Gymnasium Bayern*. Berlin: Cornelsen Verlag.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological science in the public interest*, *8*(2), 53–96.

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

Krauss, S., Martignon, L., & Hoffrage, U. (1999). Simplifying Bayesian Inference: The General Case. In N. e. a. Magnani (Ed.), *Model-based Reasoning in Scientific Discovery* (pp. 165–179), Kluwer Academic/Plenum Publishers, New York.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of einstellung. *Psychological monographs*, *54*(6), i–95.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

McCloy, R., Beaman, C. P., Morgan, B., & Speed, R. (2007). Training conditional and cumulative risk judgements: the role of frequencies, problem-structure and einstellung. *Applied Cognitive Psychology*, *21*(3), 325–344.

McDowell, M., & Jacobs, P. (2017). Meta-Analysis of the Effect of Natural Frequencies on Bayesian Reasoning. *Psychological bulletin*, *143*(12), 1273–1312.

Micallef, L., Dragicevic, P., & Fekete, J. (2012). Assessing the effect of visualizations on Bayesian reasoning through crowdsourcing. Visualization and Computer Graphics. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2536–2545.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Siegrist, M., & Keller, C. (2011). Natural frequencies and Bayesian reasoning: the impact of formal education and problem context. *Journal of Risk Research*, *14*, 1039–1055.

Staatsinstitut für Schulqualität und Bildungsforschung [ISB] (Ed.). (2016). *LehrplanPLUS Gymnasium Mathematik 10*. http://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/10/mathematik.