

Summary

How would your physician explain the accuracy of a diagnostic test he has just prescribed for you? Would he say that the test “practically never” overlooks the symptom, or “in only 0.5% of the cases” or “in 50 out of 10,000 cases?” Would any of these formulations be better or worse than the others, with respect to your comprehension of the information? The research that serves as a starting point for this dissertation suggests that the representation format of statistical information has indeed a strong impact on how this information is interpreted. More specifically, the so-called *natural frequency* format has been repeatedly shown to improve statistical thinking in different types of text problems. Natural frequencies are absolute frequencies as they result from sequentially observing and counting events in a natural environment. The goal of this dissertation is to explore whether natural frequencies can also be a useful tool in the applied domain of medical risk communication, that is, how natural frequencies can be used to improve statistical thinking in physicians and patients. *Chapter 1* introduces the concept of natural frequencies and argues that both physicians and patients need statistical thinking in order to be able to make well-informed medical decisions.

In *Chapter 2*, natural frequencies serve as a basis for a tutorial on teaching Bayesian reasoning (relevant for interpreting diagnostic test results) to medical students. The so-called representation-learning approach instructs students how to translate statistical information in terms of percentages or probabilities into a more intuitive representation, namely natural frequencies. This approach had originally been developed by Sedlmeier and Gigerenzer (2001) and implemented in a computer tutorial. The goal of Study 1 was to adapt the tutorial to the instructional setting that is still most common in the education of medical students: the traditional classroom. A one-hour classroom tutorial based on the representation learning approach was developed, and its effectiveness in this setting was compared to that of a traditional rule-learning approach. Evaluation took place two months after training by testing students’ ability to correctly solve a Bayesian inference task with information represented as probabilities. While both approaches improved performance, almost three times as many students were able to profit from the representation-training as opposed to the rule-training. Although absolute performance rates were lower in the classroom than in the computer tutorial, the difference between representation-training and rule-training had a similar magnitude. I concluded that the beneficial effect of teaching representations instead of rules appears to be robust because it can be found in even less optimal learning environments.

Previous research on natural frequencies typically used diagnostic inference problems to assess the effect of this format on statistical thinking, as did Study 1. How can we be sure that natural frequencies will also facilitate statistical thinking outside the specific context of text problems? The answer is that natural frequencies have some general features that are also likely to facilitate statistical thinking outside the text problem paradigm. One feature is that natural frequencies always specify the class of events that they refer to and thus prevent misunderstandings. A second feature would be that they are, as cardinal numbers, generally easier to understand than fractions. The goal of *Chapter 3* was to clarify whether natural frequencies have the latter feature, because there have been seemingly inconsistent results on this in the literature. Study 2 found support for the hypothesis that cardinal numbers are easier to process than fractions in diagnostic inference problems that are otherwise equivalent (i.e., even when the number of computations does not differ between the formats, as is the case in tasks with a so-called *short information menu*). Study 3 showed that the disadvantage of representing numerical information as fractions cannot be circumvented by simply adding one cardinal number, here the grand total of all considered cases. Thus, in cases where no further instructions on how to deal with percentages or probabilities can be given, a representation in terms of natural frequencies is the most promising approach to facilitating statistical insight.

The remaining chapters of the dissertation explored a second application of the tool of natural frequencies in the medical domain: How can natural frequencies be used to educate medical lay people about the uncertainties and risks associated with diagnostic tests? To explore this question, I chose one specific example of a diagnostic test: the screening mammography. In a screening mammography, women who do not show any symptoms of breast cancer get an x-ray picture of their breasts with the goal to detect breast cancer in its early stages to reduce breast cancer mortality. In order to make an informed decision about participation in mammography screening, women need to be educated about all the risks and benefits of the procedure in a manner that is both detailed and understandable. In *Chapter 4*, I analyzed how currently available German mammography pamphlets typically inform women about these issues. The descriptive analysis of 27 pamphlets revealed several problems that could cause misperceptions in the readers. First, the pamphlets were not balanced: Many relevant pieces of information about the utility, the risks, and especially the meaning of the screening results were communicated too rarely. Second, many statements were presented verbally rather than as precise statistics. Third, the highly important information about the benefit of mammography screening was often presented ambiguously, that is, without specifying the reference class. To avoid misunderstandings, it is necessary to supplement the

current health pamphlets and make the information on mammography screening more precise, more up-to-date and more transparent.

In *Chapter 5*, I designed a new pamphlet text was designed that tried to overcome the problems of the previous pamphlets. Two versions of this model pamphlet—one in which the statistical information was presented in terms of natural frequencies, and a second in which it was presented as percentages—were given to a sample of women between 40 and 69 years, that is, the main readership of mammography pamphlets. The first goal of Study 5 was to explore whether *understanding* of the model pamphlet depends on the statistical format used in the pamphlet. More specifically, it assessed how well women could recall the statistical information given in the pamphlet and how they evaluated the pamphlet in terms of understandability and relevance. The second goal was to assess the *information demand* of the women concerning mammography screening, because a pamphlet that fails to address the information demand of its audience is likely to be inefficient. The overall results showed only a small effect of the statistical format on learning from the pamphlets. Natural frequencies improved the understanding of specific topics only, for example, the benefit of mammography screening. Both model pamphlets were evaluated positively, but one third of the women indicated a dislike for precise numerical information in mammography pamphlets. The information demand of the participants was high, the women wanted to be informed especially about risks, error rates, procedural aspects, and benefits of screening. Several implications of these findings for the design of health information materials are discussed.

In concluding, I compared the effectiveness of the tool of natural frequencies in the two applications of this dissertation, namely in teaching statistical thinking to medical students and educating lay people about the risks and uncertainties involved in diagnostic tests. The effectiveness of natural frequencies appears to be influenced by moderating factors present in the specific applied context. I argue that the next step should be to take the first results concerning usability and effectiveness of the tool that I presented in this dissertation back to basic research, and to identify relevant variables that allow theoretically informed predictions about how the tool will influence statistical thinking in different (and not only applied) contexts.