

Academic Leadership Journal in Student Research

Volume 4 Spring 2016

Article 1

2016

Helping Students Understand Health Statistics

Adam J. Armijo
Fort Hays State University

Jisook Park
Fort Hays State University, j_park7@fhsu.edu

W. Trey Hill
Fort Hays State University, wthill@fhsu.edu

Follow this and additional works at: <https://scholars.fhsu.edu/aljsr>

 Part of the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Armijo, Adam J.; Park, Jisook; and Hill, W. Trey (2016) "Helping Students Understand Health Statistics," *Academic Leadership Journal in Student Research*: Vol. 4 , Article 1.
Available at: <https://scholars.fhsu.edu/aljsr/vol4/iss1/1>

This Article is brought to you for free and open access by FHSU Scholars Repository. It has been accepted for inclusion in Academic Leadership Journal in Student Research by an authorized editor of FHSU Scholars Repository.

Helping Students Understand Health Statistics

Adam J. Armijo
Fort Hays State University
Masters of Science Student
Experimental Psychology

Jisook Park
Fort Hays State University
Assistant Professor
Department of Psychology

W. Trey Hill
Fort Hays State University
Assistant Professor
Department of Psychology

Abstract

Mathematical and statistical information is often presented to patients after receiving health care. Previous research would suggest that people often have difficulty understanding statistical information, especially when it is presented in a single event probability format. According to previous research, the difficulty arises from numerical literacy, presentation format and an interaction of both. The goal of the current study is to determine the accuracy of students' estimates of having an STD, after receiving a positive test result in a simulated clinical setting. Three different formats were manipulated to help students understand the statistical information frequencies, single event probability, and an icon array. Contrary to previous research, a three step hierarchical logistic regression determined that none of the formats were aiding in the accuracy of students' estimates of having chlamydia. In fact, very few students estimated the correct likelihood they had an STD after receiving a positive test result. Possible limitations and future research are discussed.

In today's world people encounter a functionally limitless amount of information, and much of it is in the form of mathematical and statistical information (Brase & Hill, 2015; Hill & Brase, 2012). These pieces of numerical information can come from many different sources, and may reside in many different societal domains. For example, nutrition labels, health advertisements, scientific research results, and television commercials all present information in a numerical format, and often these presentations are not performed with the intent of being transparent; often, companies are trying to sell products to consumers, so the numerical information is presented in such a way that it makes the product more appealing.

Although people encounter numerical information on a daily basis, they often also interpret it incorrectly (Casscells, Schoenberger, & Graboys, 1978; Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). This problem is most apparent in health settings where physicians are trying to present health information—often in the form of statistics—to patients, with the hope that patients can become more involved in their own healthcare decisions (Gigerenzer, Gaizmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008). This contemporary model of medical decision making is known as the “shared” or “patient-oriented” model of medical decision making. In this model, where an interaction occurs between doctors and patients, it is important that both parties understand and are able to produce and understand statistical processes and terminology. It is crucial for the physicians to present information in a way that will help patients understand statistical information, so the patients will be able to make decisions in their best interest. There is also an emphasis to apply research findings to practice in order to create an evidence-based health practice (Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008). The basic skills needed to make an informed decision are reading, writing, and numeracy (Reyna, Nelson, Han, & Diekmann 2009).

Recent interest in creating affordable health care suggests there will be an influx of patients seeking out health care service. This will naturally lead to more instances of shared decision making between the physician and the patient, and these instances of shared decision making will result in an interaction centered around statistical information. Patients will then have to be able to make important decisions about what treatments to seek based on information presented in a numerical format. Informed decision will need to be made dependent on the ability to understand graphs, tables, charts, and basic statistical or mathematical information. So too, physicians will have to convey health statistical information to their patients in a numerical format.

Understanding the Difficulty of Statistical and Mathematical Information

Why do people have problems making decisions with numerical information? From the literature on medical decision making, the psychological study of judgment and decision making, there are two clear possibilities: Numerical literacy (sometimes shortened to “numeracy”), and number format (the way information is presented to people). Both these variables can create problems for patients and physicians while sitting in a doctor's office interpreting health information.

Numerical Literacy

Numerical illiteracy is a commonly noted problem in the United States and many other countries. Lipkus, Samsa, and Rimer (2001) showed about 21% of even well-educated people had difficulty in answering a simple numeracy question such as “Is 1 in 1,000 equivalent to 0.1%?” In one study,

physicians did show a better understanding of numeracy than the general public; nonetheless, 25% of the physicians in the study missed the same question (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008). Considering that physicians often have to communicate the test results with their patients, the numerical literacy of physicians, as well as patients, is of importance. For example, a test result is frequently presented as either a patient has tested positively or negatively for a certain disease. However, a simple positive or negative result can be misleading when neglecting the accuracy of the test and the base rate of the disease. The actual probability of having the disease (true positive) is not very often communicated with the patients and thus it sometimes creates great anxiety in them.

A national survey of 1,000 German citizens showed the majority of the general public believed that HIV test results were accurate with an absolute certainty (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008), though these test results may have been a false positive or a false negative result. For example, assume a hypothetical test has a 60% chance that the result will be a positive test result, and if the patient is unaffected there is still a 10% chance that the result will still be positive. The true positive results should be very clearly communicated between physicians and patients. Numerical ability is required to make an optimal decision in the real world. When making judgments about probabilities associated with prostate cancer, greater numeracy was associated with more accuracy (Hamm, Brad, & Scheid, 2003).

It is important for physicians to be able to interpret and understand health statistics. Communicating the test results effectively to their patients will enable the patients to make more informed decisions for their medical treatment. Thus, to improve statistical literacy of physicians and patients, it is critical to understand the psychological principles behind the numerical literacy. A strong supporting argument stems from the idea that the human mind is not designed to efficiently analyze and interpret medical results in certain formats (Casscells, Schoenberger, & Graboys, 1978; Cosmides & Tooby, 1996; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008; Gigerenzer & Hoffrage, 1995).

Number Format

People often have difficulty understanding probabilities when explained as percentages (e.g. 90% chance). For example, people have a hard time understanding Bayesian reasoning, which is an estimation of a probability or posterior probability, due to the frequent use of the percentage format (Cosmides & Tooby, 1996). Estimating a posterior probability is analogous to estimating the likelihood of having a disease after receiving a test result. Most test results are presented in a single event probability format. Since there is difficulty in answering questions in a single event probability, due to the percentage format, one can argue the human mind has not evolved in a way to solve a later probability (Brase & Hill, 2015). In order to solve this evolutionary problem of estimating a later probability, some argue our ancestors used frequencies over the probability format, and thus support the greater benefits of the natural frequency format (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995; Hill & Brase, 2012).

Frequencies and icon array (frequencies presented in a picture format) have been shown to accurately increase the clarity of information when making a latter judgment; thus the accuracy of Bayesian reasoning increases because of the presentation of this format (Brase, 2014; Hill &

Brase, 2012; McDowell & Jacobs, 2014). Applying this basic principle to real life, one can expect patients to estimate a later probability more accurately when a physician uses this format to communicate test results with individuals.

It is known that people with high numerate abilities tend to be able to get more out of formats than low numerates. High numerates also do similar on the formats and do not typically perform better when presented with one or the other. They are able to extract the information and use it or change it if needed. Low numerates, on the other hand, are not so talented and seem to be at the restraint of the information that is presented in front of them. However, low numerates do significantly better when presented with information in a frequency format (Peters, Vastfjall, Slovic, Mertz, Mazzocco, & Dickert, 2006). This is the interaction that is observed between numeracy and format. Only people with low numeracy actually see any significant benefit from using frequencies. People who are high in numeracy are usually able to use multiple formats and still be able to perform at a proficient level: an effect known as the fluency hypothesis in statistical reasoning (Peters et al., 2006).

This study was designed to determine the efficiency of presenting medical test results in three different formats, while also testing for numerical literacy, and presenting information sequentially in a simulated doctor's visit. Specifically, this was tested by creating a simulation of a patient's visit to a health clinic and having participants estimate their chances of having a particular disease after receiving a positive test result. Previous research would predict that the icon array format would have the most accurate posterior estimates (Brase, 2014; Garcia-Retamero & Hoffrage, 2013). It was also hypothesized that participants with low numeracy would have a greater benefit from using the icon array than participants with high numeracy.

Methods

Participants

One hundred and nineteen participants (83 females, 36 males; Mage = 19.86, SDage = 5.01; 75% Caucasian) were recruited from a midsize Midwestern University. Participants were split among the three conditions equally; natural sampled frequency (N = 39), icon array (N = 41), and single event probability (N = 39). Participants were recruited from general psychology classes for either extra credit or course credit. All data were collected following the IRB regulations.

Materials and Procedure

In a scenario, one of three formats was used to inform participants of the actual chances of having chlamydia. Formats included natural sampled frequencies (e.g., 10 out of 100) (NSF), single event probability (e.g., 10% chance) (SEP), or an icon array (IA) which was a frequency of positive and negative test results, pictographically represented (see Figure 1). One of three conditions was presented on individual computer screens. Each vignette described participants going to a health clinic for STD awareness week on campus. During the visit, participants were told they were asked to complete a demographic survey and an 11 item General Numeracy Scale (Lipkus, Samsa, & Rimer, 2001). A sample question on the general numeracy scale is "if a six sided die was rolled what would be the chances that it will come up even?" The following vignette described to

participants that they took a test for an STD, specifically chlamydia, waited a few days after taking the test and later revisited the health clinic to receive their test results. All the participants were informed of a positive test results in this hypothetical situation described in the vignette and were administered one of three formats of an explanation of actual chances of receiving a positive test result. After one of the three formats was presented, participants had to estimate their actual chances of having the STD. Then the study concluded and participants were thoroughly debriefed.

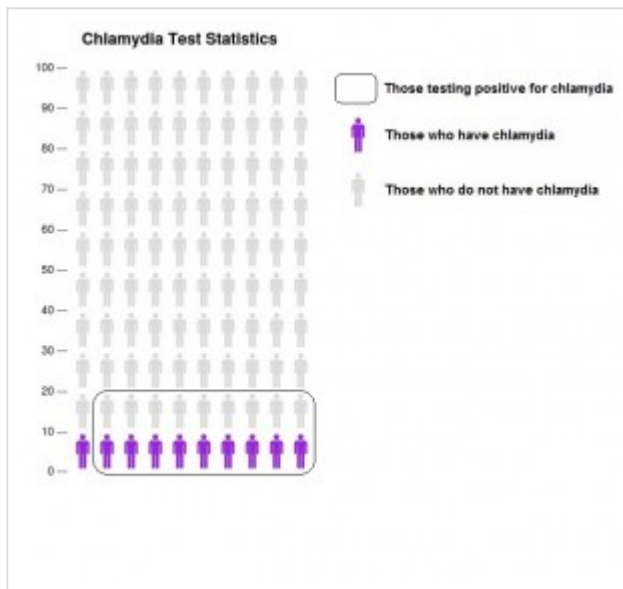


Figure 1. Icon array picture used as one manipulation in the experiment.

Results

One of the participant's data points was removed due to lack of understanding of the basic instructions of this study. A two-step hierarchical logistic regression was implemented on the participant's estimates of their likelihood of having a STD. A correct estimate was 50%, and responses within the 45-55% range were coded as correct (1), while all other responses were coded as incorrect (0). First, numerical literacy total score was entered into the model to test for predictive ability. This showed that numerical literacy was not a significant predictor of whether participants provided the correct posterior probability estimates to the STD question, Wald $\chi^2(1, N=119) < 0.01, p = .974$. Next, number format was entered into the model. This showed that number format was also not a unique predictor of correct posterior probability estimates, $\Delta\chi^2 = 1.26, p = .533$. Finally, the interaction effect of number format by numerical literacy was assessed. This product term also did not serve as a statistically significant predictor of correct or incorrect STD likelihood estimates, $\Delta\chi^2 = 0.04, p = .980$. Therefore, the current data does not support the hypotheses; a) that the icon array will have the most accurate estimation of the likelihood of having an STD after receiving a positive test result and b) participants with low numeracy will have a greater benefit from using the icon array than participants with high numeracy (Figure 2).

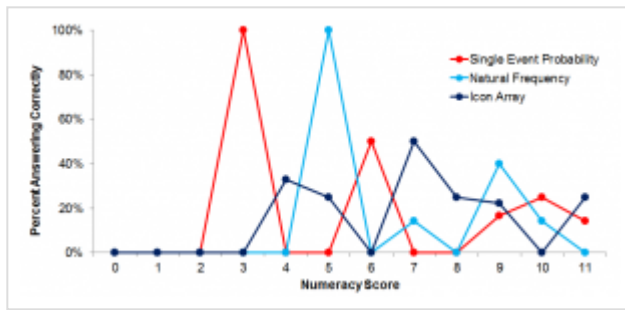


Figure 2. Relationships between number format conditions, numeracy and posterior probability estimate accuracy.

Discussion

The results of the analysis demonstrate that there was a lack of a relationship between numeracy, number format, and the accuracy of estimates. There was no support for the first hypothesis; that the icon array would lead to the most accurate estimates of having an STD. Based on previous research (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995), this is an interesting finding. Specifically, many studies in the past have shown that natural frequency and icon array number formats lead to significantly better accuracy when performing these types of Bayesian reasoning tasks.

There was also lack of support for the second hypothesis that the effect of presentation format on the accuracy of estimating likelihood of having an STD would be moderated by numeracy. Specifically, low numerates would benefit greater from the icon array when estimating the likelihood of having an STD rather than high numerates who would do well with all formats. The data suggest that neither numeracy nor presentation format increased the accuracy of correct estimates. The current study did not replicate previous research when attempting to simulate a health clinic. The current research applied the theoretical background from previous research to determine if it would help students understand health statistics.

There seemed to be a floor effect, such as the participants were not able to estimate the likelihood of having an STD. There was a bimodal distribution of the estimates, with most of them clustering either near 100% or 0%. This could suggest that the participants were confused or that they were having trouble remembering the previous format. The information had followed similar implementation as previous research, but the participants had to remember the format information. This was done to replicate how the information might be presented at a clinical setting.

The current results contradict previous research, which states that NSF and IA aid in estimating a later probability more efficiently than SEP (Brase, 2014; Hill & Brase, 2012; Gigerenzer & Hoffrage, 1995; Gigerenzer et al., 2008). Previous research findings may be obtainable if similar methods to those used by Cosmides and Tooby (1996) are utilized in the current study, such as allowing the participants to estimate with the format in front of them. The current study differed from those methods by attempting to simulate how a circumstantial visit to a health clinic may occur in real life. In doing so, the participants had to recall the format information, which was not placed in front of them, when estimating the likelihood of having an STD. Participants' difficulty in recalling the information may be attributed to limited exposure to the format materials. For example, the

participants may have only looked at the format for a few minutes and then moved on. Due to the patients' limited experience with the formatted information, the patients may not have truly understood what the results meant and could not recall a correct estimate.

Another possible factor contributing to the current studies results was the sample. The sample was obtained from a public university, whereas samples from previous research were collected from Ivy League universities. Differences have been found between Ivy League universities and public universities when performing Bayesian reasoning tasks (Brase, Fiddick, & Harries, 2006). This difference could be why many of the participants seemed to not perform as well as previous studies that have been done at Ivy League schools. Additionally, amount of compensation for the participation in the research could increase motivation to perform well when estimating frequencies and probabilities; thus, demonstrating better performance at statistical reasoning (Brase, Fiddick, & Harries, 2006). In previous research, many of the participants were compensated with cash for participating. In the current study, the participants were given course credit to participate in the study. This difference in compensation could be a reason that the previous research was not replicated.

There is also the possibility that the participants did not have the numerical ability to receive a benefit from the frequency format. This would demonstrate support of the threshold hypothesis that states a minimum level of numerical ability is needed in order to be able to correctly use NSF or IA formats (Chapman & Liu, 2009). Participants in the current study had statistically significantly lower numerical ability ($M = 7.82$) than the samples from Hill and Brase (2012) ($M = 8.47$) and Peters et al. (2006) ($M = 8.40$), but levels were comparable to those of Chapman and Liu (2009) ($M = 8.08$). Research needs to look at the frequency hypotheses and determine if there is an actual numerical ability that is needed in order to receive any benefit from the presentation format. If the threshold hypothesis is true, research needs to be done to determine what that threshold is. Currently, there is not a specific level of numerical ability that will benefit from a presentation format. Research has had trouble replicating the threshold hypothesis (Hill & Brase, 2012), but it still should be considered as a reason for the current findings.

Future research needs to answer some of the previous limitations and find ways to more effectively study numeracy and format presentation. This research is necessary in order to find the best way to communicate health statistics so patients can make informed medical decisions. More research is needed so scientists and physicians can find an optimal solution to improve medical decision making that will be pragmatic to implement.

The present study aimed to determine ways to increase student understanding of health statistics, utilizing natural sampled frequencies, single event probabilities, and icon array. No significant relationship was found between numerical literacy and correct/incorrect responses. Although these results contradict previous research, they add to the understanding of numeracy and presentation format. There still needs to be more research that will focus on how people understand health statistics and how the format will affect this understanding. This is an important area that still needs to be researched because when people are able to gain a comprehensive understanding of medical test results, then they will be able to make better decisions regarding health treatment.

References

- Brase, G. L. (2014). The power of representation and interpretation: Doubling statistical reasoning performance with icons and frequentist interpretations of ambiguous numbers. *Journal of Cognitive Psychology, 1*, 81-97. doi: 10.1080/20445911.2013.861840
- Brase, G. L., Fiddick, L., & Harries, C. (2006). Participant recruitment methods and statistical reasoning performance. *The Quarterly Journal of Experimental Psychology, 59*, 965-976. doi: 10.1080/02724980543000132
- Brase, G. L., & Hill, W. T. (2015). Good fences make for good neighbors but bad science: A review of what improves Bayesian reasoning and why. *Frontiers in Psychology, 6*(340). doi: 10.3389/fpsyg.2015.00340
- Casscells, W., Schoenberger, A., & Graboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299*(18), 999-1001.
- Chapman, B. C. & Liu, J. (2009). Numeracy, frequency, and Bayesian reasoning. *Judgment and Decision Making, 4*(1), 34-40. Retrieved from: <http://journal.sjdm.org/8708/jdm8708.pdf>
- Cosmides, L. & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on the judgment under uncertainty. *Cognition, 58*, 1-73. doi: 10.1016/0010-0277(95)00664-8
- Garcia-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improved diagnostic inferences in doctors and their patients. *Social Science & Medicine, 83*, 27-33. doi: 10.1016/j.socscimed.2013.01.034
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684-704. doi: 10.1037//0033-295x.102.4.684
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Wolshin, S. (2008). Helping doctors and patients make sense of health statistics. *Psychological science in the Public Interest, 8*, 53-96. doi: 10.1111/j.1539-6053.2008.00033.x
- Hamm, R. M., Brad, D. E., & Scheid, D. C. (2003). *Influence of numeracy upon patients prostate cancer screening outcome of numeracy upon patients prostate cancer screening outcome probability judgments*. Paper Presented at the annual meeting of the society of Judgment and Decision making, Vancouver, British Columbia, Canada.
- Hill, W. T., & Brase, G. L. (2012). When and for whom do frequencies facilitate performance? On the role of numerical literacy. *Journal of Experimental Psychology, 65*, 2342-2368. doi: 10.1080/17470218.2012.687004
- Lipkus, I. M., Samsa, G. & Rimer, B. K. (2001). General Performance on a numeracy scale among highly educated samples. *Medical Decision Making, 21*, 37-44. doi: 10.1177/0272989X0102100105