HIGH SCHOOL MATHEMATICS TEACHERS' READING OF TABLES

Adam Molnar
Oklahoma State University, USA
adam.molnar@okstate.edu

People frequently do not use numeric information as providers intended. Some lapses arise from psychological issues, but more errors (even among educated professionals) come from lack of mathematical skill. Lack of training is a cause; for instance, finding probabilities from tables appears in current US school standards, but not many earlier versions. To investigate teacher knowledge, 25 US high school mathematics teachers were interviewed on tasks related to tables and conditional probability. Although participants made mistakes, their content knowledge compared favorably to the general population. Interviewed teachers recognized common misconceptions and could offer potential suggestions to help students, but teachers acknowledged their lack of experience on the subject. Discussion includes how curriculum choices might develop students' knowledge of categorical variables.

INTRODUCTION: DECISION MAKING FOR SOCIAL POLICY

In 2015, the World Bank chose an unexpected subject for its annual development report – human decision making and development policy. Many development policies attempt to provide numeric information to people in developing countries. Unfortunately, these policies often fail because people are not able to process the information as intended. The Bank decided to pay attention to processes of mind and influences of society in order to improve the design and implementation of policies that target human behavior. The authors wrote that their report draws from findings in disciplines including "neuroscience, cognitive science, psychology, behavioral economics, sociology, political science, and anthropology." (World Bank, 2015, p. 2) The authors did not include statistics or probability in the list, an unfortunate oversight.

Later in the report, the World Bank authors provide an example involving conditional probability. The writers describe an example of confirmation bias, when people unwittingly (not deliberately) prefer evidence in line with their beliefs (Nickerson, 1998). The authors asked professional World Bank employees to make a judgment about one of two two-dimensional tables with identical numbers. One table (shown in Figure 1 on the next page) asked about the effectiveness of skin cream treating a rash, a non-controversial topic. The other table asked about minimum wage laws, a controversial topic among economists. (World Bank, 2015). The highly educated development professionals demonstrated confirmation bias. About 65% of the 1,950 participants accurately identified the new skin cream as less effective, but only 45% of participants answered correctly in the controversial minimum wage scenario (World Bank, 2015, p. 183).

This task was adapted from a 2013 study by Kahan, Peters, Dawson, and Slovic, where the authors tested 1,111 US adults with the same non-controversial skin cream version and a controversial version about crime and handguns. Overall, only 41% of surveyed adults supplied a correct answer (p. 15). As in the World Bank survey, there were political effects; people shown numbers opposite their ideology answered correctly less often.

TABLE READING IN SCHOOLS

Although the two studies identified confirmation bias, in the control condition over one-third of the professional World Bank staffers and over half of the US adult population still gave an incorrect answer. Before attempting to overcome confirmation and other psychological biases, adult citizens must be able to handle computations in non-controversial situations. The ability to read tables and compute probabilities does not ensure understanding of statistics about society, but without that ability understanding cannot occur.

In the United States, the authors who wrote the most recent attempt at national standards, the Common Core State Standards Initiative, chose to include topics related to two-way tables and probabilities. In high school, 16 standards are included in sections entitled "Conditional Probability and the Rules of Probability" and "Using Probability to Make Decisions" (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, pp. 82–83).

Two standards specifically apply to situations like the World Bank and Kahan studies. Conditional probability standard #4 states that students should "construct and interpret two-way frequency tables of data when two categories are associated with each object being classified. Use the two-way table as a sample space to decide if events are independent and to approximate conditional probabilities" (p. 82). Standard #5 asks students to "recognize and explain the concepts of conditional probability and independence in everyday language and everyday situations. *For example, compare the chance of having lung cancer if you are a smoker with the chance of being a smoker if you have lung cancer.*" (p. 82, italics as shown)

Over 40 of the 50 US states currently use Common Core standards. In many of these states, standards about probability were not part of previous guidelines; anecdotal evidence suggested that teachers were not confident about their understanding of conditional probability. Based on anecdotal evidence, I was motivated to investigate what high school mathematics teachers knew about topics in the Common Core conditional probability standards. If teachers struggle, it is very likely that their students will struggle as well. In order to assist students, teachers first must be able to properly solve problems, but they need more than content knowledge. Teachers also need pedagogical expertise, "knowledge of the strategies most likely to be fruitful in reorganizing the understanding of learners" (Shulman, 1986, pp. 9–10). Relatively few prior studies have asked high school mathematics teachers about probability, and to my knowledge, none had asked teachers about conditional probability topics. I decided to conduct task-based interviews to investigate teachers' content and pedagogical knowledge.

Medical researchers have developed a new cream for treating skin rashes. New treatments often work but sometimes make rashes worse. Even when treatments don't work, skin rashes sometimes get better and sometimes get worse on their own. As a result, it is necessary to test any new treatment in an experiment to see whether it makes the skin condition of those who use it better or worse than if they had not used it.

Researchers have conducted an experiment on patients with skin rashes. In this experiment, one group of patients used the new cream for two weeks, and a second group did not use the new cream. The total number of patients in the two groups was not the same, but this does not prevent assessment of the results.

	Rash Got Better	Rash Got Worse
Patients who DID use the new skin cream	223	75
Patients who DID NOT use the new skin cream	107	21

Is using the new cream better, the same, or worse than not using the new skin cream?

Figure 1: Rash skin cream task wording.

METHOD

Between May and July 2014, I interviewed 25 US high school mathematics teachers, 8 male and 17 female, equally divided among the states of Georgia, Pennsylvania, and South Carolina. At the time of interview, all three states used Common Core standards, although South Carolina's government implemented different standards in 2015 (South Carolina Department of Education, 2015). The participants formed a convenience sample, although I tried to gather a

balanced representation of experience with probability and statistics. The median number of college probability and statistics courses taken was 2, with a range from 0 to 5; only three participants had ever taken a stand-alone probability course. Seven teachers reported prior experience teaching high school statistics, three at the Advanced Placement (AP®) level and four at the non-AP level. This is an oversampling of experienced teachers, but 13 of the 25 participants had never taught high school statistics and also had never had professional training on probability and statistics topics.

The videotaped portion of each interview consisted of discussion of four or five tasks related to Common Core standards, and then some time in open-ended conversation. First, teachers solved each task on paper. After each task solution, we discussed the answer, other possible answers, and pedagogical strategies with students. For further information on participant demographics, methods, and all interview tasks, see my doctoral dissertation (Molnar, 2015).

The two tasks described in this paper involved two-way tables, both with realistic but not authentic data. One table was the rash skin cream task of Kahan et al. (2013) shown in Figure 1. The other table was taken from the 2010 AP® Statistics exam, Form B, Question 5 (The College Board, 2010); this table is shown in Figure 2. The questions asked participants to find the additive probability that a selected adult "is a college graduate or obtains news primarily from the internet;" to find the conditional probability that if a college graduate was selected at random, the adult obtains news primarily from the internet; and to determine if the events "college graduate" and "obtains news primarily from the internet" were independent.

An advertising agency in a large city is conducting a survey of adults to investigate whether there is an association between highest level of educational achievement and primary source for news. The company takes a random sample of 2,500 adults in the city. The results are shown in the table below.

Primary Source for News	Not High School Graduate	High School Graduate but Not College Graduate	College Graduate	Total
Newspapers	49	205	188	442
Local television	90	170	75	335
Cable television	113	496	147	756
Internet	41	401	245	687
None	77	165	38	280
Total	370	1,437	693	2,500

Figure 2: Table for conditional probability question.

RESULTS

In this section, I present teacher responses for the rash skin cream and table conditional probability questions. Although the additive probability and independence questions yielded important results (such as the fact that only 3 of the 25 teachers correctly demonstrated the two events were not independent), those other questions are more probabilistic and less related to statistics in the public domain. For each question, I present content and pedagogical results. Content results include solution paths teachers tried with errors made in the solution. Pedagogical results include potential misconceptions identified and ways teachers could help their students overcome those misconceptions.

Table Conditional Probability

To solve this question, participants needed to realize that the desired probability was conditional. The survey data had 2500 people in total. There were 693 adults who were college graduates, shown in a column total. Of these 693, 245 obtained news primarily from the internet. Selecting randomly yields a probability of 245/693 = 0.354. Only one teacher failed to see the conditioning and wrote 245/2500. Of the 24 correct responses, all but one wrote 245/693 directly;

the remaining teacher used the definition of conditional probability to write (245/2500) / (693/2500). This question had few alternate paths.

When asked about possible student misconceptions, 18 teachers mentioned failing to condition, instead getting 245/2500. Other misconceptions were mentioned much less frequently; reading comprehension and using the row total instead of the column total were next most common, with 5 mentions each. To deal with misconceptions, almost all the teachers talked about how to determine the correct denominator; as one said, "so is it out of 693 or out of 2500 is my question." Three suggestions for struggling students were made or demonstrated by at least one-third of the teachers: physically marking or covering part of the table to demonstrate the conditioning; suggesting the student reread the question, particularly the first word "If"; and directly asking the student what the denominator should be.

Rash Skin Cream

Of the 25 participants, 19 answered correctly by determining that in the experiment, the new cream was less effective than not using the new skin cream, 223/298 = 75% better against 107/128 = 84%. Of the correct answers, 11 teachers compared the percentages of patients who got better; the other 8 compared both the percentages of patients who got better and patients who got worse. A few teachers mentioned that they did the second set of computations to check their first results. Of the six incorrect answers, five were comparisons of cell counts against the grand total of 426 total patients, 223/426 against 107/426. The final teacher compared the cell counts 223 against 107 without computing any fractions.

During discussion with teachers about the solution, three noteworthy issues arose: marginal totals, sample sizes, and hypothesis testing. First, the problem statement did not include row and column totals. About half the teachers wrote marginal row totals or column totals on their paper, and a few of the teachers would have preferred that the table already have totals provided because totals would simplify the problem. Second, seven teachers made a comment about the vastly different sample sizes in the two groups, 296 with new skin cream treatment against 128 without. Three teachers qualified their conclusions because of the disparity, such as one teacher who said "I think I could make a better judgment if it was the same amount of people in both cases." Third, two teachers ventured further into inference by initially suggesting a hypothesis test, and two more asked me about potential statistical significance. Suggesting a hypothesis test changes the assumptions in the experiment, from patients as population to patients as sample. (For reference, Fisher's exact test gives a two-sided p-value of 0.0574; Pearson's unadjusted chi-square test gives a p-value of 0.0472.) During the interview, I asked teachers to consider the patients as a population and base their decision on conditional probability. Nevertheless, thinking about how people's assumptions affect their analysis might be important. In the United States, many problems (like both tasks in this paper) treat results from a sample as a population in probability problems. Professors from other countries, such as Russia, found this troubling (I. Vysotskiy, personal communication, July 28, 2016)

In the interviews, four potential student misconceptions were mentioned by at least five teachers: comparing cell counts (223 vs. 107) without computing proportions, poor reading comprehension, incorrectly forming fractions from other numbers in the table, and computation issues. To respond to comparing cell counts, most teachers suggested a reminder about the remaining information in the question – either directly pointing to patients who got worse or verbally mentioning the other group. Teachers stressed careful reading in their responses, and worried that students would struggle to find the key information. As one teacher said, students are trained to think "Is this extraneous information, is this not extraneous information, do I need to know this, is this trying to lead me somewhere?" Students without much understanding of the problem might just haphazardly pick numbers and form fractions, such as 223/107. To prevent this error, many teachers stressed context, the purpose of the numerator and denominator.

DISCUSSION

The interviewed teachers solved the conditional probability tasks pretty well. They did not prepare for the interviews, and most did not have experience teaching probability. Nevertheless, their results for the rash skin cream task (76% correct) compared favorably with World Bank

college-educated professionals (65%) and the general US adult population (41%). On other questions, the teachers compared favorably with AP® Statistics test-takers. When I discussed the results with Dan Kahan (personal communication, July 23, 2014), he concluded that "people with a critical reasoning disposition are selecting in to be high school teachers. It's heartening!" I agree. The rash skin cream and conditional probability table questions were relatively direct, requiring no probability formula. On these questions, teachers were generally able to identify potential student misconceptions and come up with legitimate arguments. The teachers had more trouble with more challenging topics such as independence and Bayes' theorem.

Several researchers have demonstrated ways to help people solve these types of problems. One issue is that table representations are very compact; although easy to publish, people often struggle to easily build problem representations from tables. For instance, Sedlmeier and Gigerenzer (2001) implemented computer-based training to have people construct problem representations instead of attempt mathematical rules; representation training had greater immediate learning and temporal stability. Martignon and Wassner (2002) demonstrated that secondary students performed better with these frequency tree representations. Spiegelhalter and Gage (2015) have written about the implementation of trees and other ideas in English schools.

Despite their relatively strong performance, participants were worried about their limitations. Of the 25 teachers, 22 expressed at least one concern about teaching probability or conditional probability. Every teacher but one had a request for assistance. About half the requests were on pedagogy, about one-third on subject matter, and the remainder on curriculum. The most frequent request was for activities for classroom use; after several teachers asked for the interview tasks all participants received copies of the problems.

The results from this modest sample do not suggest generalized conclusions. Rather, they frame ideas around teacher knowledge and suggest future lines of research, such as teaching experiments about potential responses to student misconceptions and the effects of considering the skin cream patients a sample for hypothesis testing, not a population for conditional probability. Conducting investigations that lead to teaching materials and instructor supports will eventually help students become citizens with more reasoned processes of mind.

CONFERENCE DISCUSSION

After my roundtable presentation, many people offered ideas about how to help students learn to solve problems using tables. Although it's impossible to list every commenter, I would like to acknowledge Rob Gould, Milo Schield, and session discussants Ming-Wen An and Rolf Biehler. The discussants summarized the conversation on student difficulties with a crisp phrase: "Reading a table is more than just a table."

In order to make computations, a person needs a substantial amount of previously acquired knowledge. First, the reader must comprehend the table's language. Many students are taught mathematics in a language other than their primary tongue, complicating matters. Even for students learning in their native language, complex sentences lead to errors. As a teacher said during an interview, "the difficulty I experience with statistics, the words of the problem, the way the problem is written can have a significant variant on the outcome that you get in the end." Many words have specialized meanings in the mathematical sciences. For example, the word *experiment* in the rash task represents a complex concept where patients do not affect each other and the choice to give each patient the new cream is made through a randomized mechanism. A student might never have seen a statistics experiment before, or only have seen examples with equal-sized groups. As mentioned above, groups with different sizes surprised several teachers. Alternatively, a student coming to math class from chemistry lab might be thinking about another type of experiment altogether.

After reading the words, a student needs to consider the problem situation. Each set of researchers made choices in data collection, such as what to count, how to count, and how to summarize results. To take one instance, why is the outcome in the rash table dichotomous, with the categories better and worse? What about a rash that did not change? Why wait two weeks to measure improvement? Problems exist in a human environment; analysis of any problem must include learning about the environment – a topic statistics classes sometimes avoid (Best, 2008).

During the discussion, several people proposed ideas about how table reading could become part of a learning trajectory where students would study statistical concepts primarily through categorical variables, not continuous variables. Some curriculum proposals about statistical literacy include more content about categorical variables, such as association and confounding (Schield, 2010), although the proposals have seen little uptake. An alternative might be more attractive, such as a trajectory that incorporates numeracy, literacy, and problem design. Authors with suggestions include the context-based course of MacKay (2016) and the problem-solving cycle of Wild and Pfannkuch (1999). Much debate remains on the optimal learning trajectory, but if the long-term goal is the World Bank's goal, to help people learn to make sensible decisions based on statistics about society, I believe statistical learning must include aspects of the situation that are "more than just a table."

REFERENCES

Best, J. (2008). Birds—Dead and deadly: Why numeracy needs to address social construction. *Numeracy, 1*(1). http://dx.doi.org/10.5038/1936-4660.1.1.6

The College Board. (2010). 2010 *AP Statistics Free-Response questions (Form B)*. New York, NY: Author.

Kahan, D. M., Peters, E., Dawson, E., & Slovic, P. (2013). *Motivated numeracy and enlightened self-government* (Cultural Cognition Project, Working Paper No. 116). New Haven, CT: Yale Law School. doi:10.2139/ssrn.2319992

MacKay, J. (2016). Discussion: Acquiring statistical literacy and thinking. *International Statistical Review, 84*(2), 189–194.

Martignon, L., & Wassner, C. (2002). Teaching decision making and statistical thinking with natural frequencies. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Voorburg, the Netherlands: International Statistical Institute.

Molnar, A. (2015). *High school mathematics teachers' knowledge and views of conditional probability*. Unpublished doctoral dissertation, University of Georgia.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards: Mathematics*. Washington, DC: Authors.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175–220.

Schield, M. (2010). Assessing statistical literacy: Take CARE. In Bidgood, P., Hunt, N., and Jolliffe, F. (Eds.) *Assessment Methods in Statistical Education: An International Perspective* (pp. 133 – 152). Hoboken, NJ: Wiley.

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

South Carolina Department of Education. (2015). *South Carolina College- and Career-Ready Statndards for Mathematics*. Retrieved from http://ed.sc.gov/instruction/standards-learning/mathematics/standards/

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14.

Spiegelhalter, D., & Gage, J. (2015) What can education learn from real-world communication of risk and uncertainty? *The Mathematics Enthusiast, 12*(1): Article 4.

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 221-248.

World Bank. (2015). *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: Author. doi: 10.1596/978-1-4648-0342-0.