

# Chapter 2

## Statistical Numeracy for Health\*

Mirta Galesic and Rocio Garcia-Retamero

**Abstract** Statistical numeracy is essential for understanding health-related risks and making informed medical decisions. However, this concept has not been investigated on the level of the general population or compared cross-culturally. In this chapter, we describe research that sought (1) to investigate differences in the level of statistical numeracy between two countries with different education and medical systems—the United States and Germany; (2) to study the relationship between statistical numeracy and demographic characteristics such as age, sex, and education; and (3) to test whether a subjective measure of numeracy is a valid indicator of objective measures. In a survey of probabilistic national samples in the United States and Germany, participants answered about two-thirds of the items testing objective numeracy. German participants had somewhat higher numeracy skills than participants in the USA. There was a large gap in numeracy skills between people with lower and higher educational levels, particularly in the USA. Subjective estimates of numeracy were a good indicator of the objective measures. Physicians should be aware that many patients may not understand all information relevant to making an informed decision.

---

\*In this chapter, we partially reproduced the article Galesic, M., & Garcia-Retamero, R. (2010). Statistical numeracy for health: A cross-cultural comparison with probabilistic national samples. *Archives of Internal Medicine*, 170, 462–468.

M. Galesic, Ph.D. (✉)  
Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development,  
Lentzeallee 94, Berlin 14195, Germany  
e-mail: galesic@mpib-berlin.mpg.de

R. Garcia-Retamero, Ph.D.  
Departamento de Psicología Experimental, Facultad de Psicología, University of Granada,  
Campus Universitario de Cartuja s/n, Granada 18071, Spain  
Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development,  
Lentzeallee 94, Berlin 14195, Germany  
e-mail: retamer@ugr.es

## 2.1 Introduction and Background

What percentage is 20 of 100? For most readers of this book, the answer is straightforward. Many patients, however, have difficulties grasping this and other basic statistical concepts (Davids et al. 2004; Lipkus et al. 2001; Schwartz et al. 1997; Woloshin et al. 2001). Statistical numeracy is part of a more general concept of quantitative or mathematic literacy (Golbeck et al. 2005; Speros 2005) and includes understanding the concept of a random toss and knowing how to perform elementary calculations with percentages (Davids et al. 2004; Lipkus et al. 2001). This knowledge is essential for understanding risks associated with different diseases, medical screenings, and treatments, and, consequently, for making informed decisions about health (Cokely and Kelley 2009; Estrada et al. 2004; Nelson et al. 2008; Peters et al. 2006; Reyna and Brainerd 2007; Rothman et al. 2006). This chapter describes a cross-cultural study investigating three important unanswered questions about statistical numeracy in the health context.

First, are there differences in the level of statistical numeracy between countries with different educational and medical systems—such as the USA and Germany? Several large national and international studies have included items that measure a broader concept of quantitative literacy, for example, the Programme for International Student Assessment (PISA 2003), the Trends in International Mathematics and Science Study (TIMSS; Gonzales et al. 2004), the National Assessment of Adult Literacy (NAAL; Kutner et al. 2006), and the International Adult Literacy Survey (IALS; Tuijnman 2000). Most of these studies, however, are limited to student populations and/or do not deal specifically with statistical numeracy—in particular not in the context of health. Given a stronger emphasis on mathematics and science education in the early grades in Germany compared with the USA (Rindermann 2007), it is possible that statistical numeracy is higher in Germany. However, the opposite could also be true. Because most health expenditure in the USA is privately based (55%) (see Chap. 1; World Health Organization 2012) and because patient-targeted advertising of prescription drugs is allowed, US residents may have more experience in dealing with information about medical risks, and consequently have higher statistical numeracy than the residents of Germany—where only 23% of health expenditure is privately based.

Second, what is the relationship between statistical numeracy and demographic characteristics such as age, sex, and education? To promote the ideal of informed and shared medical decision making (Barry 1999; Frosch and Kaplan 1999; Hanson 2008), it is essential to identify low-numeracy groups and to educate them in using quantitative statistical information or communicate information about health using nonquantitative formats such as visual displays and analogies (Edwards 2003; Galesic and Garcia-Retamero in press; Galesic et al. 2009; Garcia-Retamero and Galesic 2009), see Chaps. 7 and 9–11. However, all of the extant studies of statistical numeracy in health used nonprobabilistic samples of patients and students. Although informative about the numeracy skills of certain narrow groups, these studies do not allow for generalizations to any broader population. Consequently, they do not allow us to draw conclusions about the relationship between numeracy and demographic characteristics such as sex, age, and education.

Third, are objective measures of statistical numeracy equivalent to recently proposed subjective measures of this concept (Fagerlin et al. 2007)? In studies of convenience samples of patients and an Internet population, subjective measures were found to be less burdensome for the participants, at the same time approaching predictive validity of the objective measures of statistical numeracy (Zikmund-Fisher et al. 2007). Subjective measures of numeracy, however, have not yet been administered to probabilistic national samples that would enable researchers to study the relationship between objective and subjective numeracy in different demographic subgroups or to conduct cross-cultural comparisons.

To answer these questions, we conducted two studies on probabilistic national samples in the USA and Germany. This enabled us to compare statistical numeracy skills of adult population in these countries and in different sociodemographic groups within the countries.

## **2.2 Study 1: Investigating Objective Statistical Numeracy in Probabilistic National Samples**

In Study 1, conducted on probabilistic national samples in the USA and Germany, we investigated whether there are differences between the two countries in the level of objective statistical numeracy and sought to determine the relationship between numeracy and demographic characteristics.

### **2.2.1 Method**

#### **2.2.1.1 Participants**

Study 1 was conducted from July 10 through 24, 2008, on probabilistic national samples in the USA ( $n=1,009$ ) and Germany ( $n=1,001$ ), using panels of households selected through probabilistic random digit dial telephone surveys and afterward supplied with equipment that enabled them to complete computerized questionnaires. Thus, existing Internet access or lack thereof did not affect households' ability to become panel members. The panels—built and maintained by the online research panel Knowledge Networks in the USA [<http://www.knowledgenetworks.com>; 43,000 households (16% of those in the initial sample)] and the market research institute Forsa in Germany [<http://www.forsa.de>; 20,000 households (11% of those in the initial sample)]—allow for statistical inference to the general population. These panels were already used successfully in a number of studies in the areas of health, medicine, political and social sciences, economics, and public policy (Baker et al. 2003; Jacoby 2006; Lerner et al. 2003; Miller et al. 2006; Schlenger et al. 2002). Methodological studies have shown that data from such panels are comparable to the results obtained through traditional probabilistic surveys

(Chang and Krosnick 2009). The possibility of using computerized questionnaires enabled us to ask relatively complex questions involving numerical and visual information about medical treatments on a nationally representative sample.

Of the panel members who were invited to participate in the study, 54% in the USA and 52% in Germany completed the questionnaire. This is a good response rate for this survey mode (Vehovar et al. 2002). The sample structure is shown in Table 2.1. According to official statistics, the percentage of population with less education is much higher in Germany than in the USA, so we oversampled the less-educated population in the USA to ensure equivalent sample sizes of less-educated participants in both countries. To adjust for this and for minor discrepancies due to nonresponse, we used design (in the USA) and poststratification (in both countries) survey weights to bring the sample proportions in line with the population proportions. The goal of such weighting adjustments is to correct for known differences between sample and population in the hope of providing unbiased survey estimates (Bethlehem 2002; Gelman and Carlin 2002). Standard errors in all analyses were estimated using the Taylor series linearization method for estimating population characteristics from complex sample survey data, by means of commercially available software (SPSS Complex Samples procedures, SPSS version 17.0.1 (SPSS, Inc, Chicago, IL) and SUDAAN [RTI International, Research Triangle Park, North Carolina]; Siller and Tompkins 2006).

### 2.2.1.2 Stimuli and Procedure

Statistical numeracy was measured on a scale including three items developed by Schwartz et al. (1997), and six items developed by Lipkus et al. (2001), for a maximum score of 9 (see Table 2.2). The questions were translated into German by a native German speaker with excellent knowledge of English, back-translated into English by another person with equivalent language skills, and compared with the original English version. Any inconsistencies were resolved by a native German speaker and an excellent English speaker familiar with the research objectives. Finally, the English and German versions were compared and edited by a bilingual English and German speaker Chap. 15. When programming the questionnaire, special care was taken to ensure that the interface looked the same in the English and German versions. In sum, we believe that the materials in English and German were comparable. The Ethics Committee of the Max Planck Institute for Human Development approved the method used herein, and all participants consented to participation through an online consent form at the beginning of the survey.

## 2.2.2 Results

The statistical numeracy scale has satisfactory internal consistency: Cronbach alpha was 0.80 in the USA and 0.73 in Germany. Percent of correct answers to each of the

**Table 2.1** Structure of the sample of participants in Study 1 in terms of gender, age, and education<sup>a</sup>

	USA		Germany			
	Sample size (unweighted)	Sample % (weighted)	Population % <sup>b</sup>	Sample size (unweighted)	Sample % (weighted)	Population % <sup>c</sup>
Total	1,009	100%	100%	1,001	100%	100%
Gender						
Male	480	48.2	49.2	526	50.3	49.9
Female	529	51.8	50.8	475	49.7	50.1
Age						
25–39	242	33.0	35.7	270	31.4	32.5
40–54	394	39.2	38.3	410	38.9	39.9
55–69	373	27.8	26.1	321	29.7	27.7
Education <sup>d</sup>						
Less than high school	114	12.2	13.4	401	42.9	41.0
High school	599	31.8	31.2	346	31.1	31.7
Some college	145	26.1	26.0	102	11.0	11.5
College or higher	151	29.9	29.4	152	15.0	15.7

<sup>a</sup>Percentages have been rounded and may not total 100

<sup>b</sup>From Statistisches Bundesamt Deutschland, Microcensus (2007) <https://www-genesis.destatis.de/genesis/online>

<sup>c</sup>From the US Census Bureau, Current Population Survey (2008) Annual Social and Economic Supplement (<http://www.census.gov/cps/>)

<sup>d</sup>In Germany, less than high school includes participants who finished *Hauptschule*; high school, those who finished *Realschule*; and some college, those who obtained *Abitur*

**Table 2.2** Percent correct answers for each item of the numeracy scale by country in Study 1 (see also Chap. 15)

	% Correct	
	USA	Germany
1. Imagine that we flip a fair coin 1,000 times. What is your best guess about how many times the coin will come up heads in 1,000 flips? ____ times out of 1,000	73.2	72.6
2. In the Bingo Lottery, the chance of winning a \$10 prize is 1%. What is your best guess about how many people would win a \$10 prize if 1,000 people each buy a single ticket to Bingo Lottery? ____ person(s) out of 1,000	57.7	67.6
3. In the Daily Times Sweepstakes, the chance of winning a car is 1 in 1,000. What percent of tickets to Daily Times Sweepstakes win a car? ____ % of tickets	23.5	46.3
4. Imagine that we roll a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up even (2, 4, or 6)? ____ times out of 1,000	57.1	63.5
5. Which of the following numbers represents the biggest risk of getting a disease? 1 in 10, 1 in 100, or 1 in 1,000?	75.3	71.8
6. Which of the following represents the biggest risk of getting a disease? 1, 5, or 10%?	83.1	78.6
7. If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 1,000? ____ people	83.1	88.8
8. If the chance of getting a disease is 20 out of 100, this would be the same as having a ____% chance of getting the disease.	70.3	72.8
9. If Person A's chance of getting a disease is 1 in 100 in ten years, and Person B's risk is double that of A, what is B's risk?	57.3	54.5

items is presented in Table 2.2. For further analysis, we transformed the original scores ranging from 0 to 9 to a scale of 0 to 100%, indicating the percentage of the nine items that were answered correctly.

As shown in Table 2.3, German participants had higher numeracy skills than those in the USA: On average 69 vs. 65% of the items were answered correctly. This difference remains after controlling for differences in sex, age, education, and income between the two countries.

On the level of each country, sex, age, and education are all related to the numeracy score. In both countries, men had higher scores than women. Numeracy skills dropped with age ( $r=-0.12$  [-0.19, -0.05] in the USA, and  $r=-0.13$ , 95% CI [-0.20, -0.06] in Germany) and increased with education ( $r=0.50$  [0.44, 0.56] in the USA, and  $r=0.28$  [0.21, 0.35] in Germany) and income ( $r=0.32$  [0.25, 0.39] in the USA, and  $r=0.20$  [0.13, 0.27] in Germany). When we enter sex, age, education, and income together in a regression model, all four show independent effects in Germany, but in the USA only sex, education, and income explain differences in numeracy scores, while the effect of age was no longer present.

**Table 2.3** Average percentage of correctly answered items on the objective numeracy scale by country and demographic groups in Study 1

	USA ( <i>n</i> = 1,009)		Germany ( <i>n</i> = 1,001)	
	% correct	(SE)	% correct	(SE)
Overall	64.5	(1.3)	68.5	(1.1)
Gender				
Male	69.1	(2.0)	74.1	(1.6)
Female	60.2	(1.7)	62.9	(1.6)
Age				
25–39	66.5	(2.6)	72.8	(2.0)
40–54	67.5	(2.0)	68.3	(1.8)
55–69	57.9	(2.0)	64.3	(2.1)
Education <sup>a</sup>				
Less than high school	39.9	(3.1)	62.3	(1.7)
High school	56.4	(1.4)	67.3	(2.0)
Some college	64.5	(2.8)	79.2	(2.8)
College or higher	83.1	(1.8)	80.7	(2.5)
Household income <sup>b</sup>				
Lower third (up to ~\$30,000)	55.3	(2.2)	60.6	(2.3)
Middle third (~\$30–60,000)	60.8	(2.3)	70	(1.7)
Upper third (more than ~\$60,000)	76.5	(2.0)	74.1	(2.1)

<sup>a</sup>In Germany, “less than high school” includes participants who finished *Hauptschule*; “high school”—*Realschule*; and “some college”—*Abitur*

<sup>b</sup>In Germany, the categories are up to 18,000€, 18–36,000€, and more than 36,000€. To compare incomes, we used the average exchange rate for the 12 months preceding the interviews, cf. [http://www.federalreserve.gov/releases/h10/Hist/dat00\\_eu.txt](http://www.federalreserve.gov/releases/h10/Hist/dat00_eu.txt)

The inequality in numeracy skills was larger in the USA than in Germany, as reflected in the ratio between the scores in the 90th and 10th percentile of the participants ordered by their scores: This ratio was 4.5 in the USA vs. 3.0 in Germany. The inequality is visible, in particular, in average scores of people with low educational attainment vs. highly educated people in the USA: 40 vs. 83% correct, compared to 62 vs. 81% in Germany (see Table 2.3). We discuss the implications of these results in Sect. 2.4.

### 2.3 Study 2: Investigating Subjective Statistical Numeracy in Probabilistic National Samples

In Study 2, we investigated whether subjective measures of statistical numeracy (Fagerlin et al. 2007) correspond to objective measures (Lipkus et al. 2001; Schwartz et al. 1997) in general populations of the USA and Germany. If a subjective numeracy scale can differentiate between people with objectively low and high numeracy skills across different demographic groups, this would speak to its wide applicability. In addition, we tested whether the subjective perceptions of one’s numeracy are

**Table 2.4** Structure of the sample of participants in Study 2 in terms of numeracy, gender, age, and education

	USA ( <i>n</i> =238)		Germany ( <i>n</i> =260)	
	Low-numeracy group ( <i>n</i> =117)	High-numeracy group ( <i>n</i> =121)	Low-numeracy group ( <i>n</i> =127)	High-numeracy group ( <i>n</i> =133)
Mean numeracy	35.6	90.9	37.2	95.5
Female (%)	58.1	45.6	60.6	40.2
Mean age (years)	44.3	45.1	49.9	43.3
High educational levels (%) <sup>a</sup>	44.4	72.8	9.3	41.3

<sup>a</sup>Indicates some college and college or higher, as defined in Table 2.1.

dependent on the context in which they are measured, namely, before or after answering several difficult numerical questions. If the scale is sensitive to context, this would limit its applicability because the results in clinical practice would depend on patients' recent experiences with quantitative information.

### 2.3.1 Method

#### 2.3.1.1 Participants

Study 1 participants were ordered by their objective numeracy scores, and those with the highest and lowest scores were invited to participate in Study 2, conducted 3 weeks after Study 1 (August 1–15, 2008), resulting in a sample of 498 participants. Basic demographic characteristics of the sample are given in Table 2.4. This sample enables us to compare low- and high-numeracy groups within each country, as well as each of those groups between countries.

In the USA, 65.8% of all participants in Study 1 completed Study 2, and in Germany, 83.1%. The response rates among high- and low-numeracy participants were similar in both countries (i.e., it was not the case that the low-numeracy group had lower response rates). The low- and high-numeracy groups in Germany represent, respectively, approximately the bottom and top third of the population sorted by numeracy scores. Because of lower response rates in the USA, the low- and high-numeracy groups represent, respectively, approximately the bottom and top 40% of the population. Nevertheless, the average numeracy scores in both groups were still somewhat lower in the USA (Table 2.4).

#### 2.3.1.2 Stimuli and Procedure

Subjective numeracy was measured with seven of the eight items developed by Fagerlin et al. (2007; see also Zikmund-Fisher et al. 2007). The items were answered on a six-point scale, where higher values indicate higher perceived numeracy.



We excluded the item “How good are you at calculating a 15% tip?” because it is culturally specific to the USA (see Table 2.5; see Chap. 15 for the translation of the items into German). Chapter 15 lists all of the items used. The questionnaire was developed in the same way as that for Study 1. Half of the participants were randomly assigned to complete these items before a set of questions involving relatively demanding numerical calculations of risk reductions and the remaining half completed the items after answering the questions (for more details on these questions, see Garcia-Retamero and Galesic 2009; see also Chap. 10).

### 2.3.2 Results

To compare the scores on the subjective numeracy scale with the objective numeracy data, we recoded each item—originally answered on a scale of 1 to 6—to be 0 when the answer was 3 or less, or 1 when the answer was 4 or higher. Mean and standard deviation of answers to each of the items are presented in Table 2.5. For further analyses, we summed the recoded answers to the seven items and transformed the resulting scores to a scale of 0–100%, indicating the percentage of answers to the seven items that reflected high subjective numeracy.

The subjective numeracy scale has satisfactory internal consistency. The Cronbach’s alpha ranged from 0.75 to 0.87 across the two countries and groups with high vs. low objective numeracy skills. The scores on the scale were not sensitive to context: They were similar when the items were positioned before or after the tasks involving difficult calculations (average before/after difference = 2.8, 95% CI [−5.4, 11.0]); this was so for high- and low-numeracy groups in both countries.

How well does the subjective numeracy scale differentiate between participants who are very high vs. very low in terms of their objective numeracy skills (as determined in Study 1)? The average subjective numeracy scores for these two extreme groups were 38.9 (SE = 4.4) and 79.0 (SE = 2.5) in the USA, and 45.5 (SE = 3.7) and 80.0 (SE = 2.7) in Germany. These differences were stable across gender, age, education, and income groups. However, compared to the differences in objective numeracy scores between the two extreme groups ( $M = 35.6$ ,  $SE = 2.8$  vs.  $M = 90.9$ ,  $SE = 1.1$  in the USA, and  $M = 37.2$ ,  $SE = 2.0$  vs.  $M = 95.6$ ,  $SE = 0.7$  in Germany; see Table 2.4), the differences in subjective numeracy scores were smaller.

How well does the subjective numeracy scale differentiate between participants who are very high vs. very low in terms of their objective numeracy skills (as determined in study 1)? The mean (SE) subjective numeracy scores for these two extreme groups were 38.9 (4.4) and 79.0 (2.5) in the USA, and 45.5 (3.7) and 80.0 (2.7) in Germany. These differences were stable across sex, age, education, and income groups. However, compared with the differences in objective numeracy scores between the two extreme groups (mean [SE], 35.6 [2.8] vs. 90.9 [1.1] in the USA, and 37.2 [2.0] vs. 95.5 [0.7] in Germany; Table 2.4), the differences in subjective numeracy scores were smaller.

**Table 2.5** Mean ratings of items in the subjective numeracy scale by country and numeracy in Study 2

	Mean (SD)			
	USA		Germany	
	Low numeracy	High numeracy	Low numeracy	High numeracy
1. How good are you at working with fractions? Not at all good–extremely good ( <i>six-point scale</i> )	2.4 (1.2)	3.9 (1.3)	2.8 (1.2)	4.2 (1.2)
2. How good are you at working with percentages? Not at all good–extremely good ( <i>six-point scale</i> )	2.4 (1.2)	4.1 (1.3)	3.3 (1.3)	4.6 (1)
3. How good are you at figuring out how much a shirt will cost if it is 25% off? Not at all good–extremely good ( <i>six-point scale</i> )	3.5 (1.4)	4.9 (1.0)	4.1 (1.3)	5.0 (0.9)
4. When reading the newspaper, how helpful do you find tables and graphs that are parts of a story? Not at all–extremely ( <i>six-point scale</i> )	3.3 (1.4)	4.5 (1.2)	3.2 (1.5)	4.7 (1.1)
5. When people tell you the chance of something happening, do you prefer that they use words (“it rarely happens”) or numbers (“there’s a 1% chance”)? Always prefer words–always prefer numbers ( <i>six-point scale</i> )	2.6 (1.7)	4.2 (1.7)	3.0 (1.8)	4.3 (1.6)
6. When you hear a weather forecast, do you prefer predictions using percentages (e.g., “there will be a 20% chance of rain today”) or predictions using only words (e.g., “there is a small chance of rain today”)? Always prefer words–always prefer percentages ( <i>six-point scale</i> )	4.2 (1.5)	5.0 (1.2)	3.0 (1.8)	4.0 (1.7)
7. How often do you find numerical information to be useful? Never–very often ( <i>six-point scale</i> )	3.3 (1.3)	5.0 (1.1)	3.9 (1.4)	4.9 (1)

## 2.4 Discussion and Conclusions

An average citizen of the USA and Germany could answer only two-thirds of nine relatively simple items testing basic statistical numeracy skills (Table 2.3). Statistical numeracy was somewhat lower for women than for men, and it dropped slightly with age but only in Germany. Across most demographic groups, German participants

achieved somewhat higher scores than did US participants. An exception was the group with the highest education, in which US participants fared somewhat better. Differences in education systems—in particular the stronger focus on mathematics and science education in Germany from an early age (Stigler et al. 1999; Tuijnman 2000)—are likely to be the main factor underlying the differences in statistical numeracy between countries.

The inequality between people with more or less education in the USA was much larger than in Germany. Although a college-educated American could answer 83.1% of items correctly, those with less than a high school diploma could do so for only 39.9% of the items. Even for those who had a high school education the average percentage of correct answers in the USA was only 56.4%, lower than the average for German participants who had not completed a high school education (62.3%; Table 2.3).

The large differences in numeracy between persons with lower and higher educational levels have varying consequences in different medical systems. For instance—at least before the new health care reform—less educated US residents are particularly likely to be in a position to have to decide about their medical care. Although 99.7% of Germans have health insurance (see Chap. 1; see also Statistisches Bundesamt Deutschland 2011), 35% of US residents—in particular, those of lower socioeconomic status—had no health insurance or insufficient coverage (Schoen et al. 2005) and had to decide whether to pay for various medical treatments and screenings (Schoen et al. 2007). Given their low statistical numeracy, they might have had difficulty making good decisions.

The present chapter, to the best of our knowledge, describes the first study investigating statistical numeracy skills in probabilistic national samples in the USA and Germany, allowing comparison of different demographic groups within each country as well as comparison between the two countries. It also describes the first cross-cultural comparison of objective and subjective measures of statistical numeracy.

At the same time, a limitation of the studies is that levels of numeracy in the general population could be even lower than our results suggest. To become members of the national panels from which our samples were selected, participants had to accept having a computer or special TV set with Internet access installed in their homes. It is possible that people with low numeracy refused this more often than did those with high numeracy skills. On the other hand, our sample represents accurately the overall population in terms of education. Furthermore, there is no particular reason to expect that numeracy but not general educational level would be related to higher rates of refusal.

Our findings have clear implications for medical practice. Physicians should not assume that all patients can understand simple statistical indicators that are often used to express risks and benefits of medical screenings and treatments. For example, approximately 20% of the participants in both Germany and the USA could not say which of the following numbers represents the biggest risk of getting a disease: 1, 5, or 10%. Ratios were even more difficult—almost 30% could not answer whether 1 in 10, 1 in 100, or 1 in 1,000 represents the largest risk. Similarly, almost 30% of the study participants in both countries could not state what percentage 20

of 100 is, and most (76.5% and 53.7% of the participants in the USA and Germany, respectively) could not transform 1 of 1,000 to a percentage. Furthermore, many participants lacked the understanding of the concept of random toss. When asked how many times a fair coin would come up heads in 1,000 flips, more than one-fourth of the study participants in both countries gave answers that were obviously incorrect (less than 400 or more than 600 times).

Given the low levels of statistical numeracy of many patients, physicians could use items from the subjective numeracy scale to identify patients who may have problems understanding numerical information. If they have such a patient, physicians could communicate risks and benefits of treatments by means of formats that do not require high levels of numeracy, such as visual displays (see Chaps. 9, 10, and 11; see also Hanson 2008; Galesic et al. 2009; Lipkus 2007; Lipkus and Hollands 1999) and analogies (see Chap. 7; see also Garcia-Retamero and Galesic 2009; Galesic and Garcia-Retamero 2012), rather than numerical expressions. In this way, patients with low numeracy skills could understand statistical information and make better decisions about their health.

## References

- Baker, L., Wagner, T. H., Singer, S., & Bundorf, M. K. (2003). Use of the Internet and email for health care information: Results from a national survey. *Journal of the American Medical Association*, 289, 2400–2406.
- Barry, M. J. (1999). Involving patients in medical decisions: How can physicians do better? *Journal of the American Medical Association*, 282, 2356–2357.
- Bethlehem, J. G. (2002). Weighting nonresponse adjustments based on auxiliary information. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 275–288). New York: Wiley.
- Chang, L., & Krosnick, J. A. (2009). National surveys via RDD telephone interviewing versus the Internet: Comparing sample representativeness and response quality. *Public Opinion Quarterly*, 73, 641–678.
- Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, 4, 20–33.
- Davids, S. L., Schapira, M. M., McAuliffe, T. L., & Nattinger, A. B. (2004). Predictors of pessimistic breast cancer risk perceptions in a primary care population. *Journal of General Internal Medicine*, 19, 310–315.
- Edwards, A. (2003). Communicating risks through analogies. *British Medical Journal*, 327, 749.
- Estrada, C. A., Martin-Hryniewicz, M., Peek, B. T., Collins, C., & Byrd, J. C. (2004). Literacy and numeracy skills and anticoagulation control. *The American Journal of the Medical Sciences*, 328, 88–93.
- Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a Math test: Development of the subjective numeracy scale. *Medical Decision Making*, 27, 672–680.
- Frosch, D. L., & Kaplan, R. M. (1999). Shared decision making in clinical medicine: Past research and future directions. *American Journal of Preventive Medicine*, 17, 285–294.
- Galesic, M., & Garcia-Retamero, R. (2012). Using analogies to communicate information about health risks. *Applied Cognitive Psychology*. doi: 10.1002/acp.2866
- Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: overcoming low numeracy. *Health Psychology*, 28, 210–216.

- Garcia-Retamero, R., & Galesic, M. (2009). Communicating treatment risk reduction to people with low numeracy skills: A cross-cultural comparison. *American Journal of Public Health, 99*, 2196–2202.
- Gelman, A., & Carlin, J. B. (2002). Poststratification and weighting adjustments. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 289–302). New York: Wiley.
- Golbeck, A. L., Ahlers-Schmidt, C. R., Paschal, A. M., & Dismuke, S. E. (2005). A definition and operational framework for health numeracy. *American Journal of Preventive Medicine, 29*, 375–376.
- Gonzales, P., Guzmán, J. C., Partelow, L., Pahlke, E., Jocelyn, L., Kastberg, D., et al. (2004). *Highlights from the trends in international mathematics and science study (TIMSS) 2003*. U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Hanson, J. L. (2008). Shared decision making: Have we missed the obvious? *Archives of Internal Medicine, 168*, 1368–1370.
- Jacoby, W. G. (2006). Value choices and American public opinion. *American Journal of Political Science, 50*, 706–723.
- Kutner, M., Greenberg, E., Jin, Y., & Paulsen, C. (2006). *The health literacy of America's adults: results from the 2003 national assessment of adult literacy (NALS)*. U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Lerner, J. S., Gonzalez, R. M., Small, D. A., & Fischhoff, B. (2003). Effects of fear and anger on perceived risks of terrorism: A national field experiment. *Psychological Science, 14*, 144–150.
- Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. *Medical Decision Making, 27*, 696–713.
- Lipkus, I. M., & Hollands, J. G. (1999). The visual communication of risk. *Journal of the National Cancer Institute. Monographs, 25*, 149–163.
- 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.
- Miller, J. D., Scott, E. C., & Okamoto, S. (2006). Public acceptance of evolution. *Science, 313*, 765–766.
- Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine, 35*, 261–274.
- Peters, E., Vastfjall, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science, 17*, 407–413.
- Programme for International Student Assessment (PISA). (2003). *The PISA 2003 assessment framework—Mathematics, reading, science, and problem solving knowledge and skills*. Paris: OECD.
- Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences, 17*, 147–159.
- Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21*, 667–706.
- Rothman, R. L., Housam, R., Weiss, H., Davis, D., Gregory, R., Gebretsadik, T., et al. (2006). Patient understanding of food labels—The role of literacy and numeracy. *American Journal of Preventive Medicine, 31*, 391–398.
- Schlenger, W. E., Caddell, J. M., Ebert, L., Jordan, B. K., Rourke, K. M., Wilson, D., et al. (2002). Psychological reactions to terrorist attacks—Findings from the national study of Americans' reactions to September 11. *Journal of the American Medical Association, 288*, 581–588.
- Schoen, C., Doty, M. M., Collins, S. R., & Holmgren, A. L. (2005). Insured but not protected: how many adults are underinsured? *Health Affairs, W5*-289–W5-302.
- Schoen, C., Osborn, R., Doty, M. M., Bishop, M., Peugh, J., & Murukutla, N. (2007). Toward higher-performance health systems: Adults' health care experiences in seven countries, 2007. *Health Affairs, 26*, W717–W734.

- Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. (1997). The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine*, *127*, 966–972.
- Siller, A. B., & Tompkins, L. (2006). *The Big Four: Analyzing complex sample survey data using SAS, SPSS, STATA, and SUDAAN*. Paper presented at the SAS Users Group International conference (SUGI) 31, San Francisco, CA. <http://www.nesug.org/proceedings/nesug05/pos/pos3.pdf>, accessed July 30, 2012.
- Speros, C. (2005). Health literacy: Concept analysis. *Journal of Advanced Nursing*, *50*, 633–640.
- Statistisches Bundesamt Deutschland (2011). Sozialeleistungen Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus) - Fachserie 13 Reihe 1.1 - 2007. <https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/HaushalteMikrozensus/KrankenversicherungMikrozensus2130110079004.html>, accessed July 30, 2012.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States*. U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.
- Tuijnman, A. (2000). *Benchmarking adult literacy in America: An international comparative study: International adult literacy survey (IALS)*. Washington, DC: Department of Education, OVAE.
- Vehovar, V., Batagelj, Z., Lozar Manfreda, K., & Zaletel, M. (2002). Nonresponse in Web surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 229–242). New York: Wiley.
- Woloshin, S., Schwartz, L. M., Moncur, M., Gabriel, S., & Tosteson, A. N. A. (2001). Assessing values for health: Numeracy matters. *Medical Decision Making*, *21*, 382–390.
- World Health Organization (2012). *World Health Statistics 2012*. [http://www.who.int/gho/publications/world\\_health\\_statistics/2012/en/](http://www.who.int/gho/publications/world_health_statistics/2012/en/), accessed July 30, 2012.
- Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitation. *Medical Decision Making*, *27*, 663–671.