Statistical Numeracy for Health  

A Cross-cultural Comparison With Probabilistic National Samples

Mirta Galesic, PhD; Rocio Garcia-Retamero, PhD

Background: Statistical numeracy is essential for understanding health-related risks and making informed medical decisions. However, this concept has not been investigated with probabilistic national samples or compared cross-culturally. We sought (1) to investigate differences in the level of statistical numeracy between 2 countries with different educational 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.

Methods: In a survey of probabilistic, representative national samples in Germany and the United States, conducted in July and August 2008, we asked questions testing objective and subjective statistical numeracy.

Results: German participants had higher numeracy skills than did US participants. On average, 68.5% (SE, 1.1%) and 64.5% (SE, 1.3%), respectively, of items testing objective numeracy were answered correctly. Subjective estimates of numeracy were a good indicator of the objective measures. There is a large gap in numeracy skills between persons with lower and higher educational levels, particularly in the United States.

Conclusions: Physicians should be aware that many patients may not understand all information relevant to making an informed decision. Fortunately, they can identify such patients and use nonnumerical presentation formats, such as graphical displays and analogies, to communicate important statistical information.

Arch Intern Med. 2010;170(5):462-468

What percentage is 20 of 100? For most readers of this article, the answer is straightforward. Many patients, however, have difficulties grasping this and other basic statistical concepts.1-4 Statistical numeracy is part of a more general concept of quantitative or mathematical literacy5,6 and includes understanding the concept of a random toss and knowing how to perform elementary calculations with percentages.1,2 This knowledge is essential for understanding risks associated with different diseases, medical screenings, and treatments and, consequently, for making informed decisions about health.7-12 The present article describes a cross-cultural study investigating 3 important unanswered questions about statistical numeracy in the health context.

See Invited Commentary at end of article

in the context of health. Given a stronger emphasis on mathematics and science education in the early grades in Germany compared with the United States,17 it is possible that statistical numeracy is higher in Germany. However, the opposite could also be true. Because most health expenditure in the United States is privately based (55%)18 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.

See also page 470

First, are there differences in the level of statistical numeracy between countries with different educational and medical systems, such as the United States and Germany? Several large national and international studies have included items that measure a broader concept of quantitative literacy (eg, the Programme for International Student Assessment,13 Trends in International Mathematics and Science Study,14 National Assessment of Adult Literacy,15 and International Adult Literacy Survey).16 Most of these studies, however, are limited to student populations and/or do not deal specifically with statistical numeracy, in particular not...
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,\textsuperscript{19-21} 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.\textsuperscript{22-25} 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?\textsuperscript{26} 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.\textsuperscript{27} Subjective measures of numeracy, however, have not yet been administered to probabilistic, representative 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 2 studies on probabilistic national samples in the United States and Germany. This enables us to compare—for the first time, to our knowledge—statistical numeracy skills of adult populations in these countries and in different sociodemographic groups within the countries. In the first study, we investigated objective statistical numeracy skills in both countries. In the second study, we tested the applicability of a subjective numeracy scale as a proxy for objective skills.

### METHODS

The Ethics Committee of the Max Planck Institute for Human Development approved the methods used herein, and all participants consented to participation through an online consent form at the beginning of the survey.

#### STUDY 1

**Sample**

Study 1 was conducted from July 10 through 24, 2008, on probabilistic national samples in the United States (\(n=1009\)) and Germany (\(n=1001\)) using panels of households selected through probabilistic telephone (random-digit dial) 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 market research institute Forsa in Germany (http://www.forsa.de; 20 000 households [11% of those in the initial sample]) and the online research panel Knowledge Networks in the United States (http://www.knowledgenetworks.com; 43 000 households [16% 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.\textsuperscript{28-32} Methodological studies have shown that data from such panels are comparable to the results obtained through traditional probabilistic surveys.\textsuperscript{33} 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, 52.0% in Germany and 53.8% in the United States completed the questionnaire. This is a very good response rate for this survey mode.\textsuperscript{34} The sample structure is shown in Table 1 in text.

#### Table 1. Sample Structure in Study 1

<table>
<thead>
<tr>
<th></th>
<th>Germany</th>
<th>United States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unweighted Sample Size, No. of Participants</td>
<td>1001</td>
<td>1009</td>
</tr>
<tr>
<td>Weighted % of Sample</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>% of Population</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

\[a\] Percentages have been rounded and may not total 100.

\[b\] From Statistisches Bundesamt Deutschland, Microcensus, 2007 (https://www-genesis.destatis.de/genesis/online; accessed September 15, 2008).


\[d\] In Germany, less than high school includes participants who finished Hauptschule; high school, those who finished Realschule; and some college, those who obtained Abitur.
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 German and English speaker. When programming the questionnaire, special care was taken to ensure that the interface looked the same in the German and English versions. In sum, we believe that the materials in English and German were comparable.

STUDY 2

To investigate the equivalence of objective and subjective measures of numeracy skills in general populations of Germany and the United States, we conducted a study with a subset of the sample of participants in study 1.

Sample

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. Basic demographic characteristics of the sample are given in Table 3. This sample enables us to compare low- and high-numeracy groups within each country as well as each of those groups between countries.

In Germany, 83.1% of all participants in study 1 completed study 2, and in the United States, 65.8%. The response rates among high- and low-numeracy participants were similar in both countries (eg, it was not the case that the low-numeracy group had lower response rates). The low- and high-numeracy groups within each country as well as each of those groups between countries.

so we oversampled the less-educated population in the United States to ensure equivalent sample sizes of less-educated participants in both countries. This was important because the study was conducted within a project that focused specifically on less-educated people. To adjust for this and for minor discrepancies due to nonresponse, we used design (in the United States) 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. 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, Illinois] and SUDAAN [RTI International, Research Triangle Park, North Carolina]).

Table 2. Statistical Numeracy Scale

<table>
<thead>
<tr>
<th>Questions</th>
<th>Germany Answering Correctly</th>
<th>United States Answering Correctly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Imagine that we flip a fair coin 1000 times. What is your best guess about how many times the coin will come up heads in 1000 flips? ____ times out of 1000</td>
<td>72.6</td>
<td>73.2</td>
</tr>
<tr>
<td>2. In the Bingo Lottery, the chance of winning a $10 prize is 1%. What is your best guess about how many people will win a $10 prize if 1000 people each buy a single ticket for Bingo Lottery? ____ person(s) out of 1000</td>
<td>67.6</td>
<td>57.7</td>
</tr>
<tr>
<td>3. In the Daily Times Sweepstakes, the chance of winning a car is 1 in 1000. What percentage of tickets for the Daily Times Sweepstakes win a car? ____% of tickets</td>
<td>46.3</td>
<td>23.5</td>
</tr>
<tr>
<td>4. Imagine that we roll a fair, 6-sided die 1000 times. Of 1000 rolls, how many times do you think the die will come up even (2, 4, or 6)? ____ times out of 1000</td>
<td>63.5</td>
<td>57.1</td>
</tr>
<tr>
<td>5. Which of the following numbers represents the biggest risk of getting a disease? 1 in 100, 1 in 1000, or 1 in 10?</td>
<td>71.8</td>
<td>75.3</td>
</tr>
<tr>
<td>6. Which of the following represents the biggest risk of getting a disease? 1%, 10%, or 5%?</td>
<td>78.6</td>
<td>83.1</td>
</tr>
<tr>
<td>7. If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 1000? ____ person(s) out of 1000</td>
<td>88.8</td>
<td>83.1</td>
</tr>
<tr>
<td>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.</td>
<td>72.8</td>
<td>70.3</td>
</tr>
<tr>
<td>9. If person A’s chance of getting a disease is 1 in 100 in 10 years and person B’s risk is double that of A, what is B’s risk?</td>
<td>54.5</td>
<td>57.3</td>
</tr>
</tbody>
</table>

Table 3. Sample Structure of Study 2

<table>
<thead>
<tr>
<th></th>
<th>Germany (n=260)</th>
<th>United States (n=238)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-Numeracy Group (n=127)</td>
<td>High-Numeracy Group (n=133)</td>
</tr>
<tr>
<td></td>
<td>Mean numeracy</td>
<td>Female, %</td>
</tr>
<tr>
<td></td>
<td>37.2</td>
<td>60.6</td>
</tr>
<tr>
<td></td>
<td>95.5</td>
<td>40.2</td>
</tr>
</tbody>
</table>

a Indicates some college and college or higher, as defined in Table 1.
numeral counterfactuals, see Garcia-Retamero and Galesic.23 The items were answered on a 6-point scale, where higher values indicate higher perceived numeracy. Nevertheless, the average numeracy scores in both groups were still somewhat lower in the United States (Table 3).

### Questionnaire

Subjective numeracy was measured with 7 of the 8 items developed by Fagerlin et al.26,27 The items were answered on a 6-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 United States. We excluded the item “How good are you at working with percentages? 1 = not at all good; 6 = extremely good” because it is culturally specific to the United States.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Germany</th>
<th>United States</th>
</tr>
</thead>
<tbody>
<tr>
<td>How good are you at working with fractions? 1 = not at all good; 6 = extremely good</td>
<td>2.8 (1.2)</td>
<td>2.4 (1.2)</td>
</tr>
<tr>
<td>How good are you at working with percentages? 1 = not at all good; 6 = extremely good</td>
<td>3.3 (1.3)</td>
<td>2.4 (1.2)</td>
</tr>
<tr>
<td>How good are you at working with fractions? 1 = not at all good; 6 = extremely good</td>
<td>4.1 (1.3)</td>
<td>3.5 (1.4)</td>
</tr>
<tr>
<td>When reading the newspaper, how helpful do you find tables and graphs that are parts of a story? 1 = not at all helpful; 6 = extremely helpful</td>
<td>3.2 (1.5)</td>
<td>3.3 (1.4)</td>
</tr>
<tr>
<td>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.”)? 1 = always prefer words; 6 = always prefer numbers</td>
<td>3.0 (1.8)</td>
<td>2.6 (1.7)</td>
</tr>
<tr>
<td>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”)? 1 = always prefer words; 6 = always prefer percentages</td>
<td>3.0 (1.8)</td>
<td>4.2 (1.7)</td>
</tr>
<tr>
<td>How often do you find numerical information to be useful? 1 = never; 6 = very often</td>
<td>3.9 (1.4)</td>
<td>3.3 (1.3)</td>
</tr>
</tbody>
</table>

*a Adapted from Fagerlin et al.26*

As shown in Table 5, German participants had higher numeral counterfactuals than did US participants. On average, 68.5% vs 64.5% of the items were answered correctly. This difference remained after controlling for differences in sex, age, education, and income between the 2 countries.

### RESULTS

#### STUDY 1

In the first study, conducted on probabilistic national samples in the United States and Germany, we investigated whether there are differences between the 2 countries in the level of statistical numeracy and sought to determine the relationship between statistical numeracy and demographic characteristics.

The statistical numeracy scale has satisfactory internal consistency; the Cronbach α was 0.73 in Germany and 0.80 in the United States. The percentage of correct answers to each of the items is presented in Table 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 9 items that were answered correctly.

Table 4. Subjective Numeracy Scale

<table>
<thead>
<tr>
<th>Questions</th>
<th>Low-Numeracy Group Mean (SD)</th>
<th>High-Numeracy Group Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>United States</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Average Percentage of Correctly Answered Items on the Objective Statistical Numeracy Scale by Country and Demographic Groups

<table>
<thead>
<tr>
<th>Demographic Characteristic</th>
<th>Germany (n=1001)</th>
<th>United States (n=1009)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>68.5 (1.1)</td>
<td>64.5 (1.3)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>74.1 (1.6)</td>
<td>69.1 (2.0)</td>
</tr>
<tr>
<td>Female</td>
<td>62.9 (1.6)</td>
<td>60.2 (1.7)</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-39</td>
<td>72.8 (2.0)</td>
<td>66.5 (2.6)</td>
</tr>
<tr>
<td>40-54</td>
<td>68.3 (1.8)</td>
<td>67.5 (2.0)</td>
</tr>
<tr>
<td>55-69</td>
<td>64.3 (2.1)</td>
<td>57.9 (2.0)</td>
</tr>
<tr>
<td>Educational level**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>62.3 (1.7)</td>
<td>39.9 (3.1)</td>
</tr>
<tr>
<td>High school</td>
<td>67.3 (2.0)</td>
<td>56.4 (1.4)</td>
</tr>
<tr>
<td>Some college</td>
<td>79.2 (2.8)</td>
<td>64.5 (2.8)</td>
</tr>
<tr>
<td>College or higher</td>
<td>80.7 (2.5)</td>
<td>83.1 (1.8)</td>
</tr>
<tr>
<td>Household income, $**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower third (approximately $30 000)</td>
<td>60.6 (2.3)</td>
<td>55.3 (2.2)</td>
</tr>
<tr>
<td>Middle third (approximately $30 000-$60 000)</td>
<td>70.0 (1.7)</td>
<td>60.8 (2.3)</td>
</tr>
<tr>
<td>Upper third (approximately $60 000)</td>
<td>74.1 (2.1)</td>
<td>78.5 (2.0)</td>
</tr>
</tbody>
</table>

**a Educational equivalents in Germany are described in Table 1.**

**b In Germany, the categories are less than or equal to €18 000, €18 000 to €36 000, and greater than €36 000. To compare incomes, we used the average exchange rate for the 12 months preceding the interviews. (http://www.federalreserve.gov/releases/h10/Hist/dat00_eu.txt; accessed November 3, 2008).**

As shown in Table 5, German participants had higher numeracy skills than did US participants. On average, 68.5% vs 64.5% of the items were answered correctly. This difference remained after controlling for differences in sex, age, education, and income between the 2 countries.

On the level of each country, sex, age, and education were all related to the numeracy score. In both countries, men had higher scores than did women. Numeracy skills dropped with age (r = −0.13 [95% confi-
dence interval, −0.20 to −0.06] in Germany and −0.12
[−0.19 to −0.05] in the United States) and increased with
education (0.28 [0.21-0.35] in Germany and 0.50 [0.44-
0.56] in the United States) and income (0.20 [0.13-
0.27] in Germany and 0.32 [0.25-0.39] in the United
States). When we entered sex, age, education, and in-
come together in a regression model, all 4 showed inde-
pendent effects in Germany but, in the United States, only
sex, education, and income explained differences in nu-
meracy scores, whereas the effect of age was no longer
present.

The inequality in numeracy skills was larger in the
United States than in Germany, as reflected in the ratio
between the scores in the 90th and 10th percentiles of
the participants ordered by their scores. This ratio was
4.5 in the United States compared with 3.0 in Germany.
The inequality was visible in particular in the average
scores of people with low educational attainment vs highly
educated people in the United States: 39.9% vs 83.1% cor-
rect compared with 62.3% vs 80.7% in Germany (Table 5).

STUDY 2

In the second study, we investigated whether subjective
measures of statistical numeracy correspond to objec-
tive measures. If a subjective numeracy scale can differ-
entiate between people with objectively low and high nu-
meracy 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
dependent on the context in which they are measured,
namely, before or after answering several difficult nu-
merical questions. If the scale is sensitive to context, this
would limit its applicability because the results in clin-
ical practice would depend on patients’ recent experi-
ences with quantitative information.

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. Means and standard deviations of answers to each
of the items are presented in Table 4. For further analy-
ses, we summed the recoded answers to the 7 items and
transformed the resulting scores to a scale of 0% to 100%,
indicating the percentage of answers to the 7 items that
reflected high subjective numeracy.

The subjective numeracy scale has satisfactory inter-
ral consistency. The Cronbach α ranged from 0.75 to 0.87
across the 2 countries and groups with high vs low ob-
jective 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 diffi-
cult calculations (average before/after difference, 2.8 [95%
confidence interval, −5.4 to 11.0]); this was so for high-
and low-numeracy groups in both countries.

How well does the subjective numeracy scale differen-
tiate between participants who are very high vs very
low in terms of their objective numeracy skills (as de-
termined in study 1)? The mean (SE) subjective nu-
meracy scores for these 2 extreme groups were 45.5 (3.7)
and 80.0 (2.7) in Germany, and 38.9 (4.4) and 79.0 (2.5)
in the United States. These differences were stable across
sex, age, education, and income groups. However, com-
pared with the differences in objective numeracy scores
between the 2 extreme groups (mean [SE], 37.2 [2.0] vs
95.5 [0.7] in Germany and 35.6 [2.8] vs 90.9 [1.1] in
the United States; Table 3), the differences in subjective
numeracy scores were smaller.

An average citizen of Germany and the United States could
answer only two-thirds of 9 relatively simple items test-
ing basic statistical numeracy skills (Table 5). 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 partici-
pants. An exception was the group with the highest edu-
cation, in which US participants fared somewhat better.
Differences in educational systems—in particular the
stronger focus on mathematics and science education in
Germany from an early age17,38—are likely to be the main
factor underlying the differences in statistical numeracy
between countries.

The inequality between people with more or less edu-
cation in the United States was much larger than in Ger-
many. Although a college-educated American could an-
swer 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 aver-
age percentage of correct answers in the United States
was only 56.4%, lower than the average for German par-
ticipants who had not completed a high school educa-
tion (62.3%; Table 5).

The large differences in numeracy between persons
with lower and higher educational levels have varying
consequences in different medical systems. For in-
stance—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. Al-
though 99.7% of Germans have health insurance,39 35%
of US residents—in particular those of lower socioeco-
nomic status—had insufficient or no coverage40 and had
to decide whether to pay for various treatments and
screenings themselves.41 Given their low statistical
numeracy, they might have had difficulty making good
decisions.

The present article is, to the best of our knowledge,
the first to describe a study investigating statistical nu-
meracy skills in probabilistic national samples in the
United States and Germany, allowing comparison of dif-
ferent demographic groups within each country as well
as comparison between the 2 countries. It is also the first
to describe a cross-cultural comparison of objective and
subjective measures of statistical numeracy.

At the same time, a limitation of our studies is that lev-
el 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, par-
ticipants had to accept having a computer or special TV
set with Internet access installed in their homes. It is pos-
ible that people with low numeracy refused this more of-
ten 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.

This study has 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 German and US participants 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 1000 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 (53.7% of German and 76.5% of US participants) could not transform 1 of 1000 to a percentage. Furthermore, many participants lacked the understanding of the concept of random draw. When asked how many times a fair coin would come up heads in 1000 flips, more than one-fourth of the study participants in both countries gave answers that were obviously incorrect (<400 or >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,22,23,42,43 and analogies,24,25 rather than numerical expressions. In this way, patients with low numeracy skills could understand statistical information and make better decisions about their health.

Accepted for Publication: September 14, 2009.
Correspondence: Mirta Galesic, PhD, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany (galesic@mpib-berlin.mpg.de).

Author Contributions: Both authors had full access to all the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Galesic and Garcia-Retamero. Acquisition of data: Galesic and Garcia-Retamero. Analysis and interpretation of data: Galesic and Garcia-Retamero. Drafting of the manuscript: Galesic and Garcia-Retamero. Critical revision of the manuscript for important intellectual content: Galesic and Garcia-Retamero. Statistical analysis: Galesic and Garcia-Retamero. Obtained funding: Galesic and Garcia-Retamero. Administrative, technical, and material support: Galesic. Study supervision: Galesic.

Financial Disclosure: None reported.

Funding/Sponsor: This study was supported by the Foundation for Informed Medical Decision Making (United States) and the Max Planck Society (project “Helping People With Low Numeracy To Understand Medical Information”) and grant PSI2008-02019 from the Ministerio de Educación y Ciencia (project “How to Improve Understanding of Risks About Health”).

Role of the Sponsors: The funding sources did not affect the study design, data collection, analysis and interpretation of the data, writing of the report, or the decision to submit the report for publication.

References
Collective Statistical Illiteracy

The article by Galesic and Garcia-Retamero reports that a probabilistic sample of Americans and Germans could answer only two-thirds of simple statistical numeracy questions correctly. The most difficult task was to express 1 in 1000 as a percentage (question 3), which only 24% of the Americans and 46% of the Germans mastered. Furthermore, the answers reflected wider disparities between poor and rich and between less educated and higher educated respondents in the United States than in Germany. For US citizens with less than a high school education, only 40% of the questions could be answered correctly (compared with 83% for those with a college education or higher), whereas, for Germans, these percentages were only 62% and 81%, respectively. This disparity may reflect not only the stronger emphasis on math and science education in Germany but also the different attention their media pay to science. Unlike the United States, Germany has seen a German emphasis on math and science education in Germany, which only 24% of the Americans and 46% of the Germans could answer only two-thirds of simple statistical numeracy questions correctly.

What do these results tell us about medical decision making? First, one might object that the results are limited to the specific panels of households sampled by the two survey companies and to the use of nonmedical content such as the Bingo Lottery (question 2) and the Daily Times Sweepstakes (question 3). On the other hand, a national sample of 450 US adults aged 35 to 70 years that used the same basic numeracy questions but with medical content—allergic reactions from a drug—found similar results in terms of correct answers for questions 1 to 3 in Galesic and Garcia-Retamero’s statistical numeracy scale. Limited basic numeracy would not matter so much if all physicians were able to explain to patients what the percentages mean, assuming they have the time. Physicians indeed score better than the general public in basic numeracy; of 85 physicians at Dartmouth-Hitchcock Medical Center in Lebanon, New Hampshire, 100%, 91%, and 75%, respectively, correctly answered questions 1 to 3. However, the flip side is that 25% of the physicians could not correctly convert 1 in 1000 into a percentage. For both patients and physicians, understanding health statistics involves more than basic numeracy. In a European-wide study with more than 10,000 face-to-face interviews, 92% of women overestimated the cancer-specific mortality reduction of mammography screening by an order of magnitude or more, or did not know. Similarly, 89% of men overestimated the benefit of prostate-specific antigen screening. For instance, 27% of British women believed that, among 1000 women who participate in screening, 200 will be saved from dying of breast cancer. The randomized trials, in contrast, suggest a reduction from about 5 to 4 in 1000 women. One possible source for this overestimation is the common framing of this absolute reduction as a 20% relative risk reduction, which

©2010 American Medical Association. All rights reserved.