Communicating Treatment Risk Reduction to People With Low Numeracy Skills: A Cross-Cultural Comparison

Rocio Garcia-Retamero, PhD, and Mirta Galesic, PhD

Objectives. We sought to address denominator neglect (i.e. the focus on the number of treated and nontreated patients who died, without sufficiently considering the overall numbers of patients) in estimates of treatment risk reduction, and analyzed whether icon arrays aid comprehension.

Methods. We performed a survey of probabilistic, national samples in the United States and Germany in July and August of 2008. Participants received scenarios involving equally effective treatments but differing in the overall number of treated and nontreated patients. In some conditions, the number who received a treatment equaled the number who did not; in others the number was smaller or larger. Some participants received icon arrays.

Results. Participants—particularly those with low numeracy skills—showed denominator neglect in treatment risk reduction perceptions. Icon arrays were an effective method for eliminating denominator neglect. We found cross-cultural differences that are important in light of the countries' different medical systems.

Conclusions. Problems understanding numerical information often reside not in the mind but in the problem's representation. These findings suggest suitable ways to communicate quantitative medical data. (*Am J Public Health.* 2009;99: 2196–2202. doi:10.2105/AJPH.2009.160234)

Recent research on numeracy in health decision making has shown that many patients have difficulties grasping a host of numerical concepts that are necessary for them to understand health-relevant risk communications.[1–3] These communications often take the form of baseline risk estimates and risk reduction with 1 or more treatments.[4,5] Health numeracy is the individual-level skill needed to understand and use quantitative health information[6] and has a significant impact on risk perception.[7–10] Individuals with low numerical ability, for instance, are especially vulnerable to having difficulty following a complicated dosing regimen,[11] have higher history of hospitalization,[12] and are more susceptible to being influenced by the way the health information is framed in problems involving probabilities.[3]

Ratio concepts—of which risks and probabilities are examples—are particularly challenging both in medical and nonmedical contexts.[13–15] For instance, people often pay too much attention to the number of times a target event has happened and insufficient attention to the overall number of opportunities for it to happen.[16,17] This denominator-neglect effect—described by Reyna[18]—has been extensively studied.[15] In an experiment by Yamagishi,[19] participants were presented with estimates of the number of deaths in the population attributable to 11 causes of deaths (e.g., cancer) and had to assess the risk of dying of such causes. Participants rated the likelihood of a cancer killing 1286 out of 10 000 people as higher than 24.14 out of 100 people. The degree of riskiness, therefore, varied according to the number of deaths presented, irrespective of the total possible number of deaths.

Denominator neglect could have important consequences when one estimates treatment risk reduction. In medical practice, for example, the overall number of patients who receive a certain treatment is often smaller than the number of those who do not.[20–22] Similarly, patients might be able to think of more people who did not go to a certain screening or take a novel drug than they are able to think of those who did. If patients—and their physicians—disregard the overall number of treated and nontreated people (e.g., 100 and 800, respectively), they might perceive the treatment to be more effective than it actually is. Thus, they might underestimate the number of patients who died after receiving the treatment (e.g., n=5), while overestimating the number of people who died and did not receive the treatment (e.g., n=80).

To the best of our knowledge, however, most of the studies of people's perceptions of risk reductions provided samples of treated and nontreated patients of the same size,[7,23] and even experts in the field recommend doing so.[24–26] The only exception is a study conducted by Garcia-Retamero et al.[27] which showed that participants overestimated risk reduction when the overall number of treated patients was lower than the number of those who did not receive the treatment.

Yet no empirical data exist on whether people with low numeracy skills show more denominator neglect than those with high numeracy skills. As so many individuals have poor numeracy skills,[2] it is important to understand how numeracy impacts understanding of treatment risk reduction in tasks that reproduce the situations we often encounter when making health-related choices. As Fagerlin et al.[7] pointed out, low-numeracy patients might have more need for consistent denominators than would high-numeracy patients because their lack of facility with numbers puts them at a disadvantage. To our knowledge, however, this suggestion has not been investigated experimentally.

More importantly, there is a dearth of published research on whether people with low numeracy skills can be aided when making decisions about their health by using displays designed to enhance comprehension.[4,25] As Reyna and Brainerd[15] pointed out, visual displays can help people represent superordinate classes such as the overall number of patients who did and did not receive a treatment, thus reducing denominator neglect.[28] Icon arrays are graphical representations consisting of a number of circles or other icons symbolizing individuals who are affected by some risk,[25,26,29] and they have been shown to be a promising method for communicating medical risk reduction.[23,30–32] Icon arrays might then help draw people's attention to the overall number of unaffected patients and, thus, reduce denominator neglect—especially in those with low numeracy skills.

Finally, all of the studies on denominator neglect conducted so far (see Reyna and Brainerd[15] for a review) have used relatively limited laboratory samples of participants. Although these studies provide valuable information about the accuracy of understanding of these participants, because of nonprobabilistic sampling methods, the results cannot be generalized to any wider population. We conducted studies on probabilistic, representative samples in 2 countries with very different medical systems—the United States and Germany—to test the generalizability of denominator neglect and the effect of icon arrays on a wider population.

In summary, we sought to determine (1) whether participants show denominator neglect in their estimates of risk reduction, and whether participants with low numeracy show more denominator neglect than those with high numeracy; (2) whether icon arrays help reduce denominator neglect, and whether they are especially helpful for participants with low numeracy; and (3) whether US participants show more denominator neglect than German participants, and whether icon arrays improve accuracy of risk understanding in participants from both countries.

METHODS

We conducted the study in July and August of 2008, in 2 steps. We first selected probabilistic national samples of people aged 25 to 69 years in the United States (n=1009) and Germany (n=1001) by using panels of households selected through probabilistic random-digit-dial telephone surveys. The panels—built and maintained by the companies Forsa (Germany; 20 000 households; 11% of those in the initial sample) and Knowledge Networks (United States; 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 several areas, including health.[33–37] Data from such panels are comparable to the results obtained through traditional probabilistic surveys.[38,39]

Participants selected in the first step completed a numeracy scale consisting of 9 items developed by Schwartz et al.[40] and by Lipkus et al.[2] An example of an item is "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?" The scale does not contain any item that measures denominator neglect. In the second step, the participants were ordered by their numeracy scores, and those with the highest and lowest scores were invited to participate in our study, resulting in a sample of 534 participants from Germany and 513 from the United States (Table 1). This sample enabled us to compare low- and high-numeracy people within each country, as well as each of those groups between countries. In our analyses, we split the participants into 2 groups according to the median numeracy score for the total sample (i.e., 6; see Peters et al.[3] for a similar procedure). The average numeracy scores in each of the resulting groups in each country are shown in Table 1.

Response Rates

Of all panel members invited to the study, the questionnaire in step 1 was completed by 52% in Germany and 54% in the United States. Of those among them who were invited in step 2, 83% in Germany and 66% in the United States completed the questionnaire. The response rates among high- and low-numeracy participants were similar in both countries. The low- and high-numeracy groups in Germany represent approximately the bottom and top third of the population sorted by numeracy scores. Because of lower response rates in step 2 in the United States, the low- and high-numeracy groups represent approximately the bottom and top 40% of the population. Nevertheless, the average numeracy scores in both groups are still somewhat lower in the United States, reflecting the overall lower numeracy scores compared with Germany. To adjust for discrepancies due to nonresponse, we used poststratification weighting.[3]

Stimuli and Procedure

All participants completed a computerized questionnaire, which was developed in English and translated into German. The materials in English and German were back-translated and, therefore, were comparable. All translations were performed by skilled translators.

We presented participants with a medical scenario of the usefulness of Estatin—a hypothetical drug for reducing cholesterol that also decreases the risk of dying from a heart attack with a relative risk reduction of 50%. In one

TABLE 1—Sample Low- and High-Numeracy Adult Participants, by Gender and Age: United States and Germany, 2008

	Germany				United States			
	Low Numeracy (Mean = 3.39)		High Numeracy (Mean = 8.58)		Low Numeracy (Mean = 3.12)		High Numeracy (Mean = 7.82)	
	Unweighted No.	Weighted %	Unweighted No.	Weighted %	Unweighted No.	Weighted %	Unweighted No.	Weighted %
Total	257	48.6	277	51.4	237	47.8	276	52.2
Gender								
Men	108	38.5	177	61.8	89	38.1	146	52.4
Women	149	61.5	100	38.2	148	61.9	130	47.7
Age, y								
25–39	44	21.1	98	40.3	56	33.0	61	39.9
40–54	97	38.5	107	36.8	86	38.8	119	44.0
55–69	116	40.4	72	22.9	95	28.2	96	16.1

Note. Total sample for Germany, n = 534; for the United States, n = 513. Differences between low- and high-numeracy groups by gender, age, and numeracy are all larger than their respective standard errors. Our results, however, hold when we controlled for gender and age in the analyses.

condition, for instance, participants got the following information: "A new drug for reducing cholesterol, Estatin, decreases the risk of dying from a heart attack for patients with high cholesterol. Here are the results of a study of 900 such patients: 80 out of 800 of those who did not take the drug died of a heart attack, compared with 5 out of 100 of those who took the drug."

Two independent variables were manipulated between groups. Participants were randomly assigned to the conditions representing these variables. First, the overall numbers of treated and nontreated patients (i.e., sizes of the denominators) were manipulated to be 800 and 800, 100 and 800, 800 and 100, or 100 and 100, where the first and second quantity reflect the overall number of patients who did and did not take the drug, respectively (Table 2). To achieve a relative risk reduction of 50%, the sizes of the numerators (i.e., the number of treated and nontreated patients who died) varied within conditions depending on the sizes of the denominators.

Independently of these manipulations, half of the participants received—in addition to the numerical information about risk reduction—2 icon arrays presenting the risk of dying of a heart attack when the drug was and was not taken, respectively. All icon arrays contained either 800 or 100 circles depending on the overall number of patients who did and did not take the drug. Deceased patients were shown as black circles at the end of the array. An example of the condition involving icon arrays is shown in Figure 1.

As a dependent variable, we measured accuracy of risk understanding. First, following the procedure used by Schwartz et al.,[40] participants were asked how many of 1000 patients with high cholesterol might die of a heart attack if they do not take the drug. Second, they were asked how many of 1000 patients with high cholesterol might die of a heart attack it if they do take the drug. By deducting the second from the first answer and dividing it by the first, we calculated the estimated relative risk reduction.

In sum, the design of the study had 4 between-groups factors: the sizes of the denominators, icon arrays, numeracy, and nationality. To assess the effect of these factors on accuracy of risk understanding, we conducted analyses of variance (ANOVAs; see Lunney[41] and Cleary and Angel[42]). We conducted Tukey's honest significant difference test in post hoc analyses.

RESULTS

Figure 2 shows the percentage of high- and low-numeracy participants, respectively, whose estimates of risk reduction were accurate, lower, or higher than the exact value. When information about the drug was provided numerically and the sizes of the denominators were different, many participants provided inaccurate estimates. This result held especially for those participants with low numeracy. An ANOVA with numeracy and sizes of the denominators as between-groups factors on the percentage of participants whose estimates of risk reduction were inaccurate showed a main effect of numeracy ($F_{1,593}$=162.44; P=.001), and sizes of the denominator ($F_{3,593}$=16.502; P=.001), when information about the drug was provided numerically.

As Figure 2 shows, 71% of the participants with low numeracy overestimated risk reduction when the number of treated patients was lower than the number of those who did not receive the treatment (i.e., in the 100 and 800 denominator condition), whereas only 25% of the participants with high numeracy provided a lower estimate than the exact value in that condition. Note that in such a case, the number of patients who received the treatment and died (n=5) is much lower than the number of patients who did not receive the treatment and died (n=80; Table 2). Possibly, many participants—especially those with low numeracy—did not take proportions into account but only absolute numbers in the numerators, which might have led them to believe that the treatment had a larger effect than it actually did.

In contrast, 67% of the participants with low numeracy underestimated risk reduction when the number of treated patients was higher than the number of patients who did not receive treatment (i.e., in the 800 and 100 denominator condition), whereas only 19% of the participants with high numeracy provided a higher estimate than the exact value in that condition. In such a case, the number of patients who received the treatment and died (n=40) is higher than the number of patients who did not receive the treatment and died (n=10). This might have led participants—especially those with low numeracy—to believe that the treatment had a smaller effect than it actually did. Finally, when the sizes of the denominators were equal, estimated risk reduction was inaccurate in only 6% and 56% of the participants with high and low numeracy, respectively. In these conditions, participants did not necessarily have to take proportions into account to make accurate estimates but could rely on only the absolute numbers in the numerators.

Icon Arrays

As Figure 2 shows, when icon arrays were added to the numerical information, the denominator neglect effectively disappeared. Interestingly, this was particularly the case in those participants who were less skilled in using

TABLE 2—A Medical Scenario of a Hypothetical Drug for Reducing Cholesterol Showing the Number of Treated and Nontreated Patients Who Died From a Heart Attack for All Denominator Size Conditions

	Treated Patients		Nontreated Patients	
Denominator No.[a]	Deceased Patients	Population No.	Deceased Patients	Population No.
800 and 800	40	800	80	800
100 and 800	5	100	80	800
800 and 100	40	800	10	100
100 and 100	5	100	10	100

Note. Treatment risk reduction is 0.5 in all conditions.
[a]Treated and untreated people, respectively.

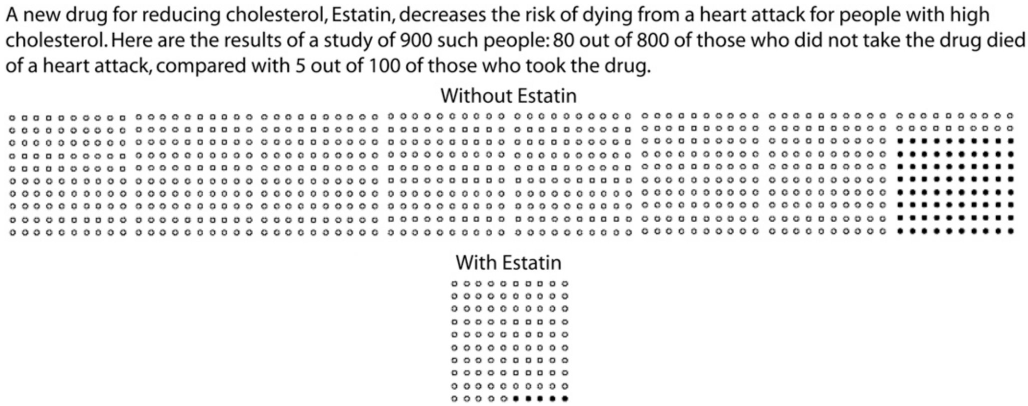

FIGURE 1—Numerical information about relative risk reduction and additional visual information (icon array).

quantitative information. An ANOVA with numeracy, sizes of the denominators, and icon arrays as between-groups factors on the percentage of participants whose estimates of risk reduction were inaccurate showed an interaction between numeracy and icon arrays ($F_{1,1100}$=6.96; P=.008) and sizes of denominators and icon arrays ($F_{3,1100}$=7.25; P=.001).

When the sizes of the denominators were different and icon arrays were added to the numerical information, the percentage of low-numeracy participants who estimated the treatment risk reduction incorrectly decreased from 74% to 42%, and from 26% to 15% in participants with high numeracy. The percentages when the sizes of the denominators were different (i.e., 42% and 15%) are similar to those when the sizes of the denominators were equal (i.e., for high numeracy, 45%; and for low numeracy, 22%). Thus, the percentages of participants who estimated risk reduction correctly were not influenced by the denominator size when icon arrays were provided. Participants, therefore, disregarded denominators when information about risk reduction was provided numerically but did not do so when icon arrays were added to the numerical information.

Comparison Between US and German Participants

Understanding medical information presented numerically was more difficult for US participants than for German participants. Icon arrays were especially useful for US participants. An ANOVA with country, sizes of the denominators, and icon arrays as between-groups factors on the percentage of participants whose estimates of risk reduction were inaccurate showed an interaction among the 3 factors ($F_{3,1100}$=3.124; P=.025).

When information about the drug was provided numerically, higher percentages of US participants (66%) provided inaccurate estimates when the sizes of the denominators were different compared with percentages of German participants (40%). When icon arrays were added to the numerical information, however, these percentages were similar in the 2 countries (31% and 36%, respectively). These percentages were also similar to the percentage of participants who provided inaccurate estimates when the sizes of the denominators were equal and icon arrays were added to the numerical information (30% and 24%, for US and German participants, respectively).

DISCUSSION

In a cross-cultural study, we addressed the effect of denominator neglect in estimates of treatment risk reduction, and analyzed whether people—especially those with low numeracy skills—can be aided by using icon arrays to enhance comprehension. Our results showed that many participants paid too much attention to numerators and insufficient attention to denominators. That is, many participants disregarded the overall number of treated and nontreated patients in favor of the number of treated and nontreated patients who died. This result held especially for those participants with low numeracy.

Our findings are compatible with previous evidence by Yamagishi,[19] who showed denominator neglect in assessments of single probabilities. Our study, however, is unique in its efforts to understand how denominator neglect is affected by numeracy. Our research is also the first to compare denominator neglect using probabilistic national samples in different countries. This is in clear contrast to previous studies, in which respondents were self-selected, preventing statistical inference to broader populations.

The result that people—especially those with low numeracy skills—could disregard crucial information when making important decisions about their health is a troubling finding. We show, however, an effective method to eliminate denominator neglect: providing icon arrays in addition to numerical information drew participants' attention to the denominators and helped them make more accurate assessments. These results support and extend our own and others' findings about the usefulness of visual aids in communicating medical risks.[7,23,26,31,32] These findings also extend the literature on denominator neglect as they provide experimental support of Reyna and Brainerd's[15] hypothesis that visual displays can help people represent superordinate classes (i.e., the overall number of patients who did and did not receive a treatment; see also Ancker et al.[25]). Finally, our results have implications for medical practice as they suggest suitable ways to communicate quantitative medical data to

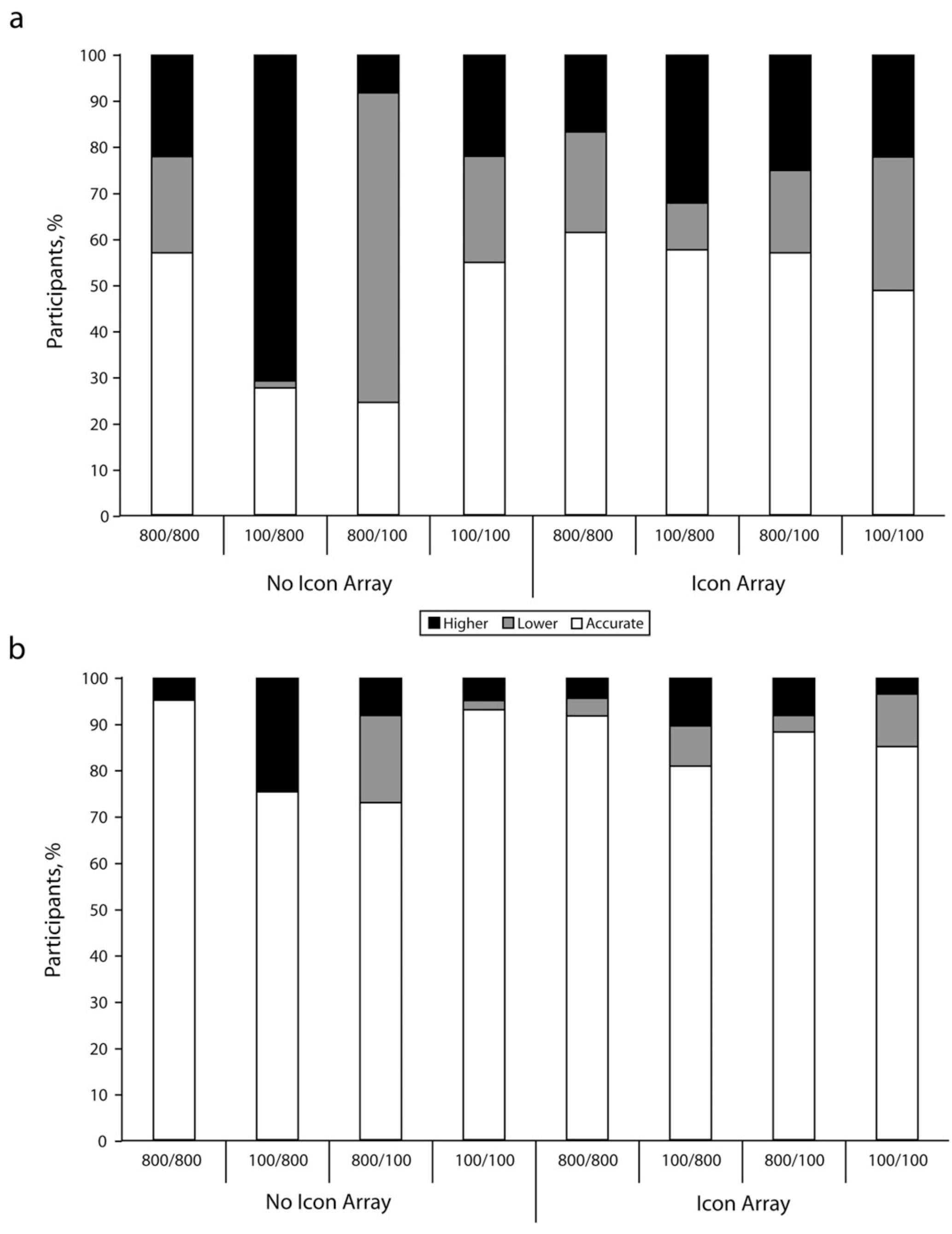

FIGURE 2—Percentage of participants whose estimates of risk reduction were either accurate or lower or higher than the exact value as a function of the sizes of the numerators and denominators and presence or absence of icon arrays, by (a) low numeracy or (b) high numeracy: United States and Germany, 2008.

people who are disadvantaged by their lack of numerical skills.

We also found interesting cross-cultural differences between participants from Germany and the United States. In line with results in national studies investigating numerical skills,[1,43,44] US participants were particularly vulnerable to having difficulty when estimating treatment risk reduction. This result can be attributable to the differences in the educational systems between Germany and the United States. In fact, the stronger emphasis on math and science education in early grades in Germany compared with the United States[45] is one of the reasons why German students score higher than US students on measures of quantitative literacy.[1,43,44] These differences might also affect people's ability to reason about numerical concepts—including risks—in adulthood. This hypothesis is supported by the fact that residents in Germany also have higher scores on numeracy than those in the United States.[3] Our findings are also compatible with research showing that US residents show more risk aversion than those in Germany.[46,47] It is possible that these cultural differences in orientation toward risk and uncertainty could have made Americans particularly vulnerable to errors when estimating treatment risk reduction.

Cross-cultural differences in accuracy of estimates of treatment risk reduction might be important in light of the different medical systems in the 2 countries. For example, most health expenditure in the United States is financed by the individual (55%),[48] and a significant part of the population either does not have health insurance (26%) or has sporadic or insufficient coverage (9%).[49] In comparison, in Germany only 23% of health expenditure is made by the individual, and most people have health insurance (99.7%).[50] This means that US citizens may be more likely than their German counterparts to be required to determine which medical treatment, if any, they need. In addition, patient-targeted advertising of pharmaceutical products is allowed in the United States but not in Germany, adding to the pressure on American patients to make their own decisions about their treatments. Our results show that American patients can be especially aided when making decisions about their treatments by using visual displays designed to enhance comprehension.

The strengths of our study are an informative between-groups design, a large sample size, and careful execution of the same study on probabilistic, national samples in 2 countries. A limitation of our research, however, is that icon arrays were added to the numerical information about treatment risk reduction that participants received in all conditions. A second presentation of the same information might have reinforced understanding of risk reduction—regardless of the information format. Previous research by Galesic et al,[23] however, showed that icon arrays are effective even when no additional numerical information is provided, supporting our conclusions about the usefulness of these methods for communicating medical risks. Another limitation of our study might be the relatively low response rate (although this rate is very good for this type of survey[51]). However, because we used probability sampling in all stages of selection, we were able to control for any sociodemographic differences between respondents and nonrespondents.

Our findings support the notion that problems in communicating risks occur because inappropriate information formats are often used and not because of biases in people's minds.[52,53] Similar reductions in what superficially looked like biased thinking were observed in the case of conditional probabilities,[54] relative risk reductions,[55] and single-event probabilities.[56] In the same vein, we show that denominator neglect in estimations about treatment risk reduction disappears when both the numerator and the denominator are presented in a transparent way. An alternative explanation of our results could be that icon arrays help people to engage in rational (i.e., numerical) processing, whereas those who received risks in less transparent ways might rely more on intuitive thought processes.[13,57,58] Whether this is in fact the case should be explored in future research. ■

About the Authors

Rocio Garcia-Retamero is with University of Granada, Granada, Spain, and the Max Planck Institute for Human Development, Berlin, Germany. Mirta Galesic is with the Max Planck Institute for Human Development.

Correspondence should be sent to Rocio Garcia-Retamero, PhD, Facultad de Psicología, Universidad de Granada, Campus Universitario de Cartuja s/n, 18071 Granada, Spain (e-mail: rretamer@ugr.es; galesic@mpib-berlin.mpg.de).

This article was accepted June 26, 2009.

Contributors

Both R. Garcia-Retamero and M. Galesic conceptualized the study, obtained funding, and wrote the article. M. Galesic led the data collection with assistance from R. Garcia-Retamero. R. Garcia-Retamero was involved in the data analysis and M. Galesic provided advice on it.

Acknowledgments

This study is part of 2 projects, "Helping people with low numeracy to understand medical information," funded by the Foundation for Informed Medical Decision Making (United States) and the Max Planck Society (Germany), and "How to improve understanding of risks about health (PSI2008-02019)," funded by the Ministerio de Educación y Ciencia (Spain).

We thank those who participated in interviews or provided feedback about the design of the study or data analyses.

Note. The authors declare independence from the funding agencies.

Human Participant Protection

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

References

1. Kutner M, Greenberg E, Jin Y, Paulsen C. *The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy (NCES 2006-483). US Department of Education.* Washington, DC: National Center for Education Statistics; 2006.

2. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. *Med Decis Making.* 2001;21:37–44.

3. Galesic M, Garcia-Retamero R. Statistical numeracy for health: a cross-cultural comparison with probabilistic national samples. *Arch Intern Med.* In press.

4. Montori VM, Rothman RL. Weakness in numbers. The challenge of numeracy in healthcare. *J Gen Intern Med.* 2005;20:1071–1072.

5. Garcia-Retamero R, Galesic M. Heuristics. In: Kattan MW, ed. *The Encyclopedia of Medical Decision Making.* Thousand Oaks, CA: SAGE Publications. 2009:56–59.

6. Ancker JS, Kaufman D. Rethinking health numeracy: a multidisciplinary literature review. *J Am Med Inform Assoc.* 2007;14:713–721.

7. Fagerlin A, Ubel PA, Smith DM, Zikmund-Fisher BJ. Making numbers matter: present and future research in risk communication. *Am J Health Behav.* 2007;31(Suppl 1):S47–S56.

8. Baker DW, Gazmararian JA, Williams MV, et al. Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *Am J Public Health.* 2002;92:1278–1283.

9. Baker DW, Parker RM, Williams MV, Clark WS, Nurss J. The relationship of patient reading ability to self-reported health and use of health services. *Am J Public Health.* 1997;87:1027–1030.

10. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. How making a risk estimate can change the feel of that risk: shifting attitudes toward breast cancer risk in a general public survey. *Patient Educ Couns.* 2005;57:294–299.

11. Estrada CA, Martin-Hryniewicz M, Peek BT, Collins C, Byrd JC. Literacy and numeracy skills and anticoagulation control. *Am J Med Sci.* 2004;328:88–93.

12. Apter AJ, Cheng J, Small D, et al. Asthma numeracy skill and health literacy. *J Asthma.* 2006;43:705–710.

13. Kahneman D. A perspective on judgment and choice: mapping bounded rationality. *Am Psychol.* 2003;58:697–720.

14. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science.* 1974;185:1124–1131.

15. Reyna VF, Brainerd CJ. Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learn Individ Differ.* 2008;18:89–107.

16. Denes-Raj V, Epstein S, Cole J. The generality of the ratio-bias phenomenon. *Pers Soc Psychol Bull.* 1995; 21:1083–1092.

17. Reyna VF. How people make decisions that involve risk: a dual-processes approach. *Curr Dir Psychol Sci.* 2004;13:60–66.

18. Reyna VF. Class inclusion, the conjunction fallacy, and other cognitive illusions. *Dev Rev.* 1991;11:317–336.

19. Yamagishi K. When a 12.86% mortality is more dangerous than 24.14%: implications for risk communication. *Appl Cogn Psychol.* 1997;11:495–506.

20. Grossarth-Maticek R, Ziegler R. Randomized and non-randomized prospective controlled cohort studies in matched pair design for the long-term therapy of corpus uteri cancer patients with a mistletoe preparation. *Eur J Med Res.* 2008;13:107–120.

21. Lichtenberg P, Levinson D, Sharshevsky Y, Feldman D, Lachman M. Clinical case management of revolving door patients—a semi-randomized study. *Acta Psychiatr Scand.* 2008;117:449–454.

22. Walitza S, Werner B, Romanos M, Warnke A, Gerlach M, Stopper H. Does methylphenidate cause a cytogenetic effect in children with attention deficit hyperactivity disorder? *Environ Health Perspect.* 2007; 115:936–940.

23. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. *Health Psychol.* 2009;28:210–216.

24. International Patient Decision Aid Standards. Criteria for Judging the Quality of Patient Decision Aids. 2005. Available at: http://www.ipdas.ohri.ca/IPDAS_checklist.pdf. Accessed September 5, 2008.

25. Ancker JS, Senathirajah Y, Kukafka R, Starren JB. Design features of graphs in health risk communication: a systematic review. *J Am Med Inform Assoc.* 2006;13: 608–618.

26. Paling J. Strategies to help patients understand risks. *BMJ.* 2003;327:745–748.

27. Garcia-Retamero R, Galesic M, Gigerenzer G. Do icon arrays help reduce denominator neglect? *Med Decision Making.* In press.

28. Lloyd FJ, Reyna VF. A Web exercise in evidence-based medicine using cognitive theory. *J Gen Intern Med.* 2001;16:94–99.

29. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. *BMJ.* 2002;324:827–830.

30. Fagerlin A, Wang C, Ubel PA. Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? *Med Decis Making.* 2005;25:398–405.

31. Stone ER, Sieck WR, Bull BE, Yates JF, Parks SC, Rush CJ. Foreground:background salience: explaining the effects of graphical displays on risk avoidance. *Organ Behav Hum Decis Process.* 2003;90:19–36.

32. Zikmund-Fisher BJ, Ubel PA, Smith DM, et al. Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. *Patient Educ Couns.* 2008;73:209–214.

33. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and email for health care information: results from a national survey. *JAMA.* 2003;289:2400–2406.

34. Jacoby WG, Sniderman PM. The structure of value choices in the American public. Paper presented at: the 2006 Annual Meeting of the Southern Political Science Association; January 5, 2006; Atlanta, GA.

35. Lerner JS, Gonzalez RM, Small DA, Fischoff B. Effects of fear and anger on perceived risks of terrorism: a national field experiment. *Psychol Sci.* 2003;14:144–150.

36. Miller JD, Scott EC, Okamoto S. Public acceptance of evolution. *Science.* 2006;313:765–766.

37. Schlenger WE, Caddell JM, Ebert L, et al. Psychological reactions to terrorist attacks: findings from the national study of Americans' reactions to September 11. *JAMA.* 2002;288:581–588.

38. Krosnick J, Nie N, Rivers D. Web survey methodologies: a comparison of survey accuracy. Paper presented at: the Annual Conference of the American Association for Public Opinion Research; May 14, 2005; Miami Beach, FL. Available at: http://communication.stanford.edu/faculty/Krosnick/2005%20AAPOR%20Presentation.ppt. Accessed October 3, 2008.

39. Krosnick J, Nie N, Rivers D. Comparing major survey firms in terms of survey satisficing: telephone and Internet data collection. Paper presented at: the Annual Conference of the American Association for Public Opinion Research; May 14, 2005; Miami Beach, FL.

40. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. *Ann Intern Med.* 1997;127: 966–972.

41. Lunney GH. Using analysis of variance with a dichotomous dependent variable - empirical study. *J Educ Meas.* 1970;7:263–269.

42. Cleary PD, Angel R. The analysis of relationships involving dichotomous dependent-variables. *J Health Soc Behav.* 1984;25:334–348.

43. Programme for International Student Assessment. *The PISA 2003 Assessment Framework—Mathematics, Reading, Science, and Problem Solving Knowledge and Skills.* Paris, France: Organisation for Economic Cooperation and Development; 2003.

44. Tuijnman A. *Benchmarking Adult Literacy in America: An International Comparative Study (International Adult Literacy Survey).* Washington, DC: Department of Education, Office of Vocational and Adult Education; 2000.

45. National Center for Education Statistics. *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. Research and Development Report.* Washington, DC: US Department of Education Office of Educational Research and Improvement; 1999. Available at: http://nces.ed.gov/pubs99/1999074.pdf. Accessed September 29, 2008.

46. Weber EU, Hsee CK. Cross-cultural differences in risk perception, but cross-cultural similarities in attitude towards perceived risk. *Manage Sci.* 1998;44:1205–1217.

47. Renn O, Rohrmann B. Cross-cultural risk perception: state and challenges. In: Renn O, Rohrmann B, eds. *Cross-Cultural Risk Perception: A Survey of Empirical Studies.* Dordrecht, The Netherlands: Kluwer; 2000.

48. World Health Organization. Core health indicators, 2008. Available at: http://www.who.int/whosis/database/core/core_select_process.cfm?countries=all&indicators=nha. Accessed January 8, 2009.

49. Schoen C, Doty MM, Collins SR, Holmgren AL. Insured but not protected: how many adults are underinsured? *Health Aff (Millwood).* 2005;(Suppl Web exclusives):W5-289–W5-302.

50. Statistisches Bundesamt Deutschland. Microcensus, 2007. Available at: http://www.genesis.destatis.de/genesis/online. Accessed November 5, 2008.

51. Vehovar V, Batagelj Z, Lozar Manfreda K, Zaletel M. Nonresponse in Web surveys. In: Groves RM, Dillman DA, Eltinge JL, Little RJA, eds. *Survey Nonresponse.* New York, NY: John Wiley & Sons; 2002:229–242.

52. Garcia-Retamero R, Galesic M. Trust in health care. In: Kattan MW, ed. *The Encyclopedia of Medical Decision Making.* Thousand Oaks, CA: SAGE Publications. 2009:41–43.

53. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest.* 2008;8:53–96.

54. Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: frequency formats. *Psychol Rev.* 1995;102:684–704.

55. Covey J. Meta-analysis of the effects of presenting treatment benefits in different formats. *Med Decis Making.* 2007;27:638–654.

56. Gigerenzer G, Hertwig R, van den Broek E, Fasolo B, Katsikopoulos K. "A 30% chance of rain tomorrow:" how does the public understand probabilistic weather forecasts? *Risk Anal.* 2005;25:623–629.

57. Sloman SA. The empirical case for two systems of reasoning. *Psychol Bull.* 1996;119:3–22.

58. Stanovich KE, West RF. Individual differences in reasoning: implications for the rationality debate. *Behav Brain Sci.* 2000;23:645–726.