Individual Differences in Graph Literacy: Overcoming Denominator Neglect in Risk Comprehension

Article *in* Journal of Behavioral Decision Making · October 2012

DOI: 10.1002/bdm.751

Yasmina Okan
University of Leeds

Antonio Maldonado López
University of Granada

Journal of Behavioral Decision Making, *J. Behav. Dec. Making* (2011)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/bdm.751

Individual Differences in Graph Literacy: Overcoming Denominator Neglect in Risk Comprehension

YASMINA OKAN[1*], ROCIO GARCIA-RETAMERO[1,2], EDWARD T. COKELY[3,2] and ANTONIO MALDONADO[1]

[1]*Department of Experimental Psychology, University of Granada, Granada, Spain*

[2]*Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany*

[3]*Department of Cognitive and Learning Sciences, Michigan Technological University, Houghton, MI, USA*

ABSTRACT

Graph literacy is an often neglected skill that influences decision making performance. We conducted an experiment to investigate whether individual differences in graph literacy affect the extent to which people benefit from visual aids (icon arrays) designed to reduce a common judgment bias (i.e., denominator neglect—a focus on numerators in ratios while neglecting denominators). Results indicated that icon arrays more often increased risk comprehension accuracy and confidence among participants with high graph literacy as compared with those with low graph literacy. Results held regardless of how the health message was framed (chances of dying versus chances of surviving). Findings contribute to our understanding of the ways in which individual differences in cognitive abilities interact with the comprehension of different risk representation formats. Theoretical, methodological, and prescriptive implications of the results are discussed (e.g., the effective communication of quantitative medical data). Copyright © 2011 John Wiley & Sons, Ltd.

KEY WORDS individual differences; denominator neglect; graph literacy; visual aids; risk perception; risk communication

The popular saying "a picture is worth a thousand words" reflects the widespread belief that pictures and graphical displays can facilitate the communication and comprehension of complicated information. In modern societies, there is a growing need for such simplification. For instance, research has documented many ways in which doctors and patients struggle to grasp numerical concepts that are prerequisites for the accurate evaluation and communication of risks (Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007; Peters et al., 2006; Schwartz, Woloshin, Black, & Welch, 1997). Fortunately, tools such as visual displays—including line plots or bar charts—can help overcome some of these difficulties in professionals and the public alike (Ancker, Senathirajah, Kukafka, & Starren, 2006; Fuller, Dudley, & Blacktop, 2002; Lipkus, 2007; Lipkus & Hollands, 1999). However, graphs are not equally useful for all individuals (Ancker et al., 2006; Garcia-Retamero & Galesic, 2010b; Lipkus, 2007). Recent research has shown that people differ substantially in their ability to understand graphically presented information, or *graph literacy* (Galesic & Garcia-Retamero, 2011). In this paper, we address the question of how individual differences in graph literacy influence the efficacy of visual displays.

Individuals with high graph literacy have been found to make more elaborate inferences when viewing graphical displays as compared with less graph-literate individuals. For instance, highly graph-literate individuals extract information of a higher level of complexity when viewing line graphs than do individuals with low graph literacy. They are also more likely to direct their attention to typical line graph information (e.g., quantitative trend information; Maichle, 1994). When viewing bar graphs, individuals with high graph literacy are more capable of making main effect inferences on the basis of the data represented than are less graph-literate individuals (Shah & Freedman, 2011). Moreover, novice graph viewers often neglect the relevance of important elements of graphs (Mazur & Hickam, 1993) and interpret graphs incorrectly as compared with experienced graph viewers (Shah & Hoeffner, 2002). Differences in individual-level skills such as numeracy (i.e., the ability to process basic probability and numerical concepts; Cokely, Galesic, Schulz, & Garcia-Retamero, 2011; Cokely, Ghazal, Garcia-Retamero, & Galesic, in press; Galesic & Garcia-Retamero, 2010; Peters et al., 2006; Reyna, Nelson, Han, & Dieckmann, 2009) also affect people's reactions to different graphic representation formats (Wright, Whitwell, Takeichi, Hankins, & Marteau, 2009), the extent to which visual aids are useful in the assessment of treatment risk reduction (Galesic, Garcia-Retamero, & Gigerenzer, 2009; Garcia-Retamero & Galesic, 2010b), and recall of numerical information (Garcia-Retamero & Galesic, 2011). Other studies have documented a range of other factors, such as age (Garcia-Retamero, Galesic, & Gigerenzer, 2010) or people's familiarity with the specific content depicted (Shah, 2001; Shah & Hoeffner, 2002), that tend to be related to graph literacy. For example, familiarity can affect the ease with which

*Correspondence to: Yasmina Okan, Facultad de Psicología, Universidad de Granada, Campus Universitario de Cartuja s/n, 18071 Granada, Spain.
E-mail: yokan@ugr.es

This study is part of the projects, "How to Improve Understanding of Risks about Health (PSI2008-02019)," "Helping Doctors and Their Patients Make Decisions about Health," (PSI2011-22954) and "Decision Making and Risk Behavior" (PSI2009-12217), funded by the Ministerio de Ciencia e Innovación (Spain). The authors declare independence from this funding agency in each of the following: design of the study; collection, management, analysis, and interpretation of the data; and preparation of the manuscript.

different individuals make inferences based on graphs, even among expert scientists (Roth & Bowen, 2003).

Although a growing body of evidence has documented differences in people's ability to understand graphs, relatively less research has examined how graph literacy influences the efficacy of risk communication interventions. Visual displays such as line plots or bar charts facilitate the communication of information by enabling the representation of quantitative information in spatial locations. Thus, visual displays facilitate inferences about conceptual relations in the data to be made on the basis of spatial relations (Gattis, 2002, 2004; Gattis & Holyoak, 1996; Kosslyn, 2006; Tversky, 2001). However, it is unclear how, when, and why differences in graph literacy affect the efficacy of risk communication tools such as visual aids. When people are faced with such tools, graph literacy could interact with fundamental skills needed for competent decision making and reasoning (e.g., consistency in risk perception; Bruine de Bruin, Parker, & Fischhoff, 2007; Parker & Fischhoff, 2005). Here, we report on an investigation of the extent to which individual differences in graph literacy influence the effectiveness of graphs designed to reduce a common bias in judgment and decision making, namely denominator neglect.

Denominator neglect refers to people's tendency to pay too much attention to numerators in ratios (i.e., the number of times a target event has happened) and insufficient attention to denominators (i.e., the overall opportunities for it to happen; Denes-Raj, Epstein, & Cole, 1995; Reyna, 2004; Reyna & Brainerd, 2008). This tendency has been observed in the health domain in numerous studies, leading people, for example, to judge cancer as riskier when it is described as killing 1286 of 10 000 people than as killing 24.14 of 100 people (Yamagishi, 1997), or to judge 36 500 people dying of cancer every year as riskier than 100 dying every day (Bonner & Newell, 2008). An example of denominator neglect in a medical context would be focusing on the number of treated and non-treated patients who die, without considering the overall number of treated and non-treated patients, in judging whether a treatment was effective. Consequently, people often assess treatment risk reduction inaccurately (Garcia-Retamero & Galesic, 2009).

Denominator neglect can be particularly problematic when people are required to judge the effectiveness of a treatment using information from unequally sized groups of treated and non-treated patients, which is a common situation in medical practice (e.g., Grossarth-Maticek, & Ziegler, 2008; Lichtenberg, Levinson, Sharshevsky, Feldman, & Lachman, 2008). In particular, in a loss frame (i.e., number of patients who died), when the overall number of patients who receive a treatment is smaller than the number of those who do not receive it (e.g., 100 and 800, respectively), people tend to overestimate risk reduction (Garcia-Retamero & Dhami, 2011; Garcia-Retamero & Galesic, 2009; Garcia-Retamero et al., 2010). In this situation, people would overestimate treatment risk reduction because they would take into account the absolute numbers of treated and non-treated patients who die (e.g., 2 and 80, respectively) rather than the proportion of treated and non-treated patients who die (e.g., 2 of 100 and 80 of 800, respectively).

Icon arrays (i.e., graphical representations consisting of a number of circles or other icons symbolizing individuals who are affected by some risk; Ancker et al., 2006; Edwards, Elwyn, & Mulley, 2002; Paling, 2003) are an effective method for eliminating denominator neglect and increasing the accuracy of people's risk estimates (Garcia-Retamero & Galesic, 2009; Garcia-Retamero et al., 2010). It has been suggested that icon arrays improve the accuracy of quantitative reasoning by disentangling classes that are overlapping in ratios, making part-to-whole relations visually available (e.g., Reyna, 1991; Reyna & Brainerd, 2008; see also Ancker et al., 2006). However, in line with the literature reviewed examining the comprehension of visual displays such as line plots or bar charts (e.g., Maichle, 1994; Shah & Freedman, 2011), it is possible that a certain level of graph literacy is required to associate the visual patterns contained in icon arrays with meaningful interpretations of the data represented (i.e., risk reduction information). This would imply that individual differences in graph literacy could affect the effectiveness of icon arrays in improving the accuracy of quantitative reasoning. Investigating this issue was the main aim of this paper.

We tested the hypothesis (H_1) that the effectiveness of icon arrays in reducing denominator neglect would be larger for individuals with high graph literacy than for individuals with low graph literacy. We further hypothesized (H_2) that highly graph-literate participants would report more confidence in their estimates when icon arrays are provided as compared with less graph-literate participants. We measured individual differences in graph literacy using a scale developed by Galesic and Garcia-Retamero (2011). This scale covers four frequently used graph types—i.e., line plots, bar charts, pies, and icon arrays—and has been designed to measure graph comprehension in the medical domain.

Another factor that plays an important role in risk comprehension is the structure and the content of the message (i.e., *message framing*). Previous research has demonstrated that the presentation of information in a negative versus positive frame can have a large impact on judgment and decision making (Edwards, Elwyn, Covey, Matthews, & Pill, 2001; Garcia-Retamero and Cokely, in press; Garcia-Retamero & Galesic, 2010a; Levin, Schneider, & Gaeth, 1998; Rothman & Salovey, 1997). For instance, studies in medical contexts have shown that the likelihood for people to engage in illness-detecting behaviors is larger when messages are framed in terms of potential losses, whereas gain-framed messages are more likely to lead to prevention behaviors (Banks et al., 1995; Gerend & Shepherd, 2007; Rivers, Salovey, Pizarro, Pizarro, & Schneider, 2005; Rothman, Martino, Bedell, Detweiler, & Salovey, 1999; Toll et al., 2010). Other studies have documented how manipulating the way in which risks associated with different treatments are framed (i.e., chances of surviving versus chances of dying) affects evaluations and preferences for these treatments (Haward, Murphy, & Lorenz, 2008; Marteau, 1989; McNeil, Pauker, Sox, & Tversky, 1982; Wilson, Kaplan, & Schneiderman, 1987).

In studies investigating the effect of denominator neglect on perceptions of treatment risk reduction, information has always been framed in negative terms (Garcia-Retamero & Galesic, 2009; Garcia-Retamero et al., 2010). To the best of

DOI: 10.1002/bdm

our knowledge, no study has yet analyzed whether denominator neglect equally affects the accuracy of people's estimates of risk reduction when information is presented in either negative or positive terms (i.e., chances of dying and surviving, respectively). This issue is of interest as research has demonstrated the generalized existence of an attentional bias toward negative information (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). This bias has been documented in studies analyzing people's visual search of faces (Hansen and Hansen, 1988; Öhman, Lundqvist, & Esteves, 2001), color-naming latencies in the emotional Stroop task (Pratto & John, 1991), and event-related brain potentials associated to negative versus positive stimuli (Smith, Cacioppo, Larsen, & Chartrand, 2003). A common finding in these studies is that negative stimuli elicit attention more automatically than positive stimuli. Therefore, in tasks investigating understanding of treatment risk reduction, the presentation of information in negative terms (i.e., chances of dying) rather than positive terms (i.e., chances of surviving) could exacerbate the effect of denominator neglect. The attentional bias for negative information might lead people to focus on the absolute numbers in the numerators (i.e., the number of people that die) and overlook the denominator. Investigating this issue was an additional aim of this paper. In particular, we aimed to determine whether an attentional bias for negative information could amplify the effect of denominator neglect. If this is the case, denominator neglect should be larger when information is presented in terms of chances of dying, as compared with when it is presented in terms of chances of surviving.

To test the hypotheses stated so far, we conducted a study where we analyzed participants' understanding of medical risk reductions after treatment. Participants with different levels of graph literacy were presented with scenarios involving equally effective treatments but differing in the overall number of treated and non-treated patients. In some conditions, the number of patients who did receive a treatment was equal to those who did not; in other conditions, it was smaller or larger. Some participants were provided with icon arrays alongside numerical information about risk reduction, whereas others received numerical information only. To test the hypotheses that differences in graph comprehension skills affect both the effectiveness of icon arrays in reducing denominator neglect (H_1) and people's confidence in their estimates (H_2), we compared the accuracy of risk reduction estimates and the self-reported confidence in these estimates in participants with high versus low graph literacy scores. Additionally, to analyze whether message framing could affect risk understanding, we provided half of the participants with the information for the medical scenarios in terms of chances of dying, whereas the other half received the information in terms of chances of surviving.

METHOD

Participants

Participants were 168 undergraduate students from the University of Granada, Spain (16% women, median age of 20 years, range 18–28 years). They were recruited through the University's online recruitment pool and through advertisements made during lectures, and participated in exchange of course credit. A paper-and-pencil questionnaire was completed by participants in group sessions ranging from 2 to 12 participants. The sessions were always conducted under the supervision of one of the researchers, in order to ensure that questionnaires were completed individually. The tasks relevant for the present study took between 15 and 20 minutes to complete. Afterward, participants completed other unrelated tasks for an additional 40 minutes. Participants were randomly assigned to the different experimental groups.

Stimuli and procedure

Participants were presented with four medical scenarios describing the usefulness of hypothetical new drugs for reducing cholesterol that also decreased the risk of dying from a heart attack. The order of the four scenarios was randomized. Participants read and evaluated information about the risks and subsequently completed a graph literacy scale.

Measurement of graph literacy

Graph literacy scores were collected using the instrument developed by Galesic and Garcia-Retamero (2011). This scale consists of 13 items and includes items reflecting three levels of graphical comprehension traditionally outlined in the literature (see Friel, Curcio, & Bright, 2001): (i) the ability to *read the data*, that is, to find specific information in the graph, which corresponds to the more elementary level (for instance, the ability to read off the height of a particular bar within a bar chart, or the number of icons of a particular type in an icon array); (ii) the ability to *read between the data*, that is, to find relationships in the data as shown on the graph, which corresponds to an intermediate level (for instance, the ability to read off the difference between two bars or sets of icons); and (iii) the ability to *read beyond the data*, or make inferences and predictions from the data, which corresponds to an advanced level (for example, the ability to project a future trend from a line chart, or to understand the importance of attending to scale ranges and scale labels when comparing two charts). The scale contains four items assessing the ability to *read the data*, four items assessing the ability to *read between the data*, and five items assessing the ability to *read beyond the data* (for examples of items, see Figure 1).

Additionally, the scale is designed to cover four frequently used graph types—line plots, bar charts, pies, and icon arrays—and includes items dealing with the communication of medical risks, treatment efficiency, and prevalence of diseases. In sum, the scale measures both basic graph-reading skills and more advanced graph comprehension, for different types of graphs. The psychometric properties of this scale have been assessed in a survey conducted on probabilistically representative national samples of people from Germany and the USA, demonstrating satisfactory levels of internal consistency (Cronbach's alpha of .74 in Germany and .79 in the

DOI: 10.1002/bdm

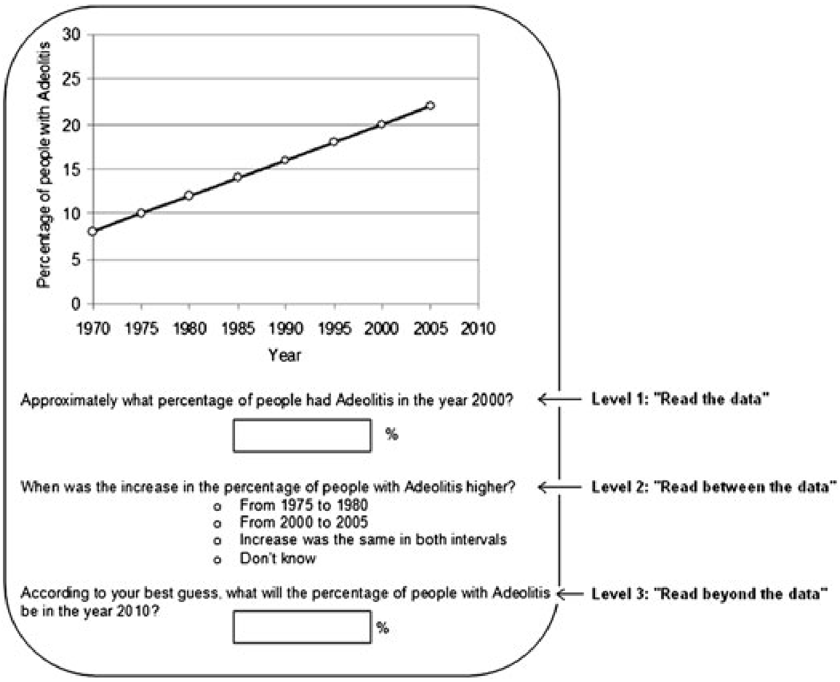

Figure 1. Examples of items measuring the three abilities of graph comprehension M. Galesic & R. Garcia-Retamero, Medical Decision Making, 31, 444–457, Copyright © 2011 by Society for Medical Decision Making. Reprinted by Permission of SAGE Publications.

USA; .70 in the sample of the current study) and convergent validity (the average correlation of the total score with graph comprehension items from existing literacy questionnaires was .44; for further details on the psychometric properties of the scale see Galesic & Garcia-Retamero, 2011).

We split participants into two groups according to the median graph literacy score for the total sample (i.e., 10). The group of participants with low graph literacy included those who obtained nine or fewer correct responses (n=68), whereas the group of participants with high graph literacy included those who obtained 10 or more correct responses (n=100). Participants with low graph literacy answered on average 7.8 items correctly (SD=2.0), whereas participants with high graph literacy answered on average 10.9 items correctly (SD=0.8).

Information about medical risks

An example of the information presented in the medical scenarios for the negative framed message condition is as follows (the original material was in Spanish).

"A new drug that reduces cholesterol, Benofreno, *decreases* the chances of *dying* after a heart attack for people with high cholesterol. Here are the results of a study of 900 people with high cholesterol: 80 out of 800 people who *did not take* the drug *died* after a heart attack, compared to 2 out of 100 people who *took* the drug."

In the positive framed message condition, the information was presented as follows.

"A new drug that reduces cholesterol, Benofreno, *increases* the chances of *surviving* aftera heart attack for people with high cholesterol. Here are the results of a study of 900 people with high cholesterol: 10 out of 100 people who *took* the drug *survived* after a heart attack, compared to 16 out of 800 people who *did not take* the drug."

The rest of the drugs were named Cenofreno, Denofreno, and Genofreno, respectively.

Design

Three independent variables were manipulated in the study. First, *message frame* was manipulated between subjects by providing half of the participants with the information for all medical scenarios in terms of chances of dying, whereas the other half received the information in terms of chances of surviving. Second, the overall numbers of treated and non-treated patients (i.e., *the sizes of the denominators*) were manipulated within subjects and were set to be either 800–800, 100–800, 800–100, 100–100, where the first and second quantities reflect the overall number of patients who did and did not take the drug, respectively. The sizes of the numerators—i.e., the number of treated and non-treated patients who died (survived) in the negative (positive) framed message condition—varied within conditions depending on the size of the denominator. In particular, the treatment always had an 80% relative risk reduction or increase in survival rate (i.e., from 10% to 2% in terms of chances of dying, or from 2% to 10% in terms of chances of surviving; see Table 1).

Finally, the presentation of *visual aids* was manipulated between subjects by providing half of the participants with two icon arrays in addition to the numerical information for each medical scenario. These icon arrays presented the risk of dying of a heart attack/surviving after a heart attack when the drug was and was not taken. All icon arrays contained either 800 or 100 circles depending on the overall number

Table 1. Number of treated and non-treated patients who die (negative framing; top panel) or survive (positive framing; bottom panel) after a heart attack for all denominator sizes

	Treated patients		Non-treated patients	
Denominator sizes	Patients who died	Population size	Patients who died	Population size
800–800	16	800	80	800
800–100	16	800	10	100
100–800	2	100	80	800
100–100	2	100	10	100
	Treated patients		Non-treated patients	
Denominator sizes	Patients who survived	Population size	Patients who survived	Population size
800–800	80	800	16	800
800–100	80	800	2	100
100–800	10	100	16	800
100–100	10	100	2	100

Note: Risk reduction/increase in survival rate is 80% in all conditions.

of patients who did and did not take the drug. The patients who died (in the negative framed message condition) or survived (in the positive framed message condition) were represented with black circles at the end of the array (see Figure 2, for an example; original material was in Spanish).

As dependent variables, we measured participants' accuracy of risk understanding after they read the information provided for each medical scenario, and their subjective confidence in the estimates given. In order to measure accuracy of risk understanding, we followed the procedure used by Schwartz et al. (1997). First, participants were asked how many of 1000 patients with high cholesterol might die of/survive after a heart attack (for negative and

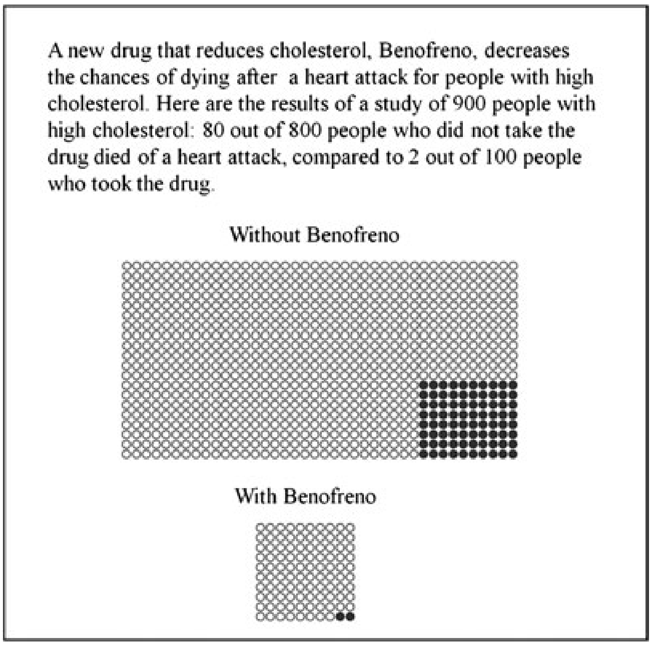

Figure 2. Numerical information about risk reduction and icon arrays that participants received. This example presents the information provided in the negative framing, 800–100 condition

positive message frame, respectively) if they do not take the drug. Second, they were asked how many of 1000 patients with high cholesterol might die of/survive after a heart attack if they did take the drug (for negative and positive message frame, respectively). By subtracting the second from the first answer and dividing it by the first, we calculated the estimated relative risk reduction for the negative message frame. For the positive message frame, we subtracted the first from the second answer and then divided it by the second, in order to calculate the estimated increase in survival rate. Participants were then classified depending on whether their estimates were accurate, lower, or higher than the exact value (i.e., 80%). Estimates were treated as correct only when they were exactly right. Participants' degree of confidence in their estimates was measured on a scale of 1–10, were 1 represented "not at all confident" and 10 represented "very confident."

In sum, our experimental design can be summarized as a 2 (high versus low graph literacy; between subjects) $\times$ 2 (positive versus negative message frame; between subjects) $\times$ 4 (size of denominators 800–800, 100–800, 800–100, 100–100; within subjects) $\times$ 2 (absence versus presence of icon arrays; between subjects) factorial design. To assess the effect of these factors on risk perceptions and confidence, we conducted analyses of variance (ANOVAs). Note that for the ANOVAs on risk perceptions that will be described in the next paragraphs, we only considered whether scores were accurate (1) or inaccurate (0). Thus, the dependent variable accuracy of risk understanding is dichotomous. We used the Bonferroni correction for post hoc analyses.

RESULTS

First, we aimed to determine whether participants showed denominator neglect in their estimates of treatment risk reduction, and to analyze differences in denominator neglect as a function of message frame. To this end, we conducted a 2×4 ANOVA with message frame as between-subjects factor and sizes of denominators as a within-subjects factor on the percentage of participants whose estimates of risk reduction were accurate in the numerical condition only (i.e., when participants did not receive icon arrays). We followed Lunney (1970; see also Cleary & Angel, 1984), who showed that ANOVAs can be used to obtain conservative results for large samples of a dichotomous dependent variable. The analysis only revealed a main effect of sizes of the denominators, $F(3,216)=6.45$, $p=.001$, $\eta_p^2=0.082$. Thus, consistent with previous research, when no icons were provided and the sizes of the denominators were equal (i.e., in the 800–800 and 100–100 conditions), participants' estimates were significantly more accurate than when the denominators were different (i.e., 800–100 and 100–800 conditions). The ANOVA did not reveal a main effect of message frame, or an interaction involving message frame implying that the percentage of accurate estimates did not reliably vary as a function of this variable ($F<1$). Thus, we did not find support for the notion that an attentional bias might exacerbate

DOI: 10.1002/bdm

the effect of denominator neglect when information is framed negatively.[1]

These results suggest that participants tended to pay too much attention to numerators and insufficient attention to denominators in both message framing conditions (see Figure 3). Accordingly, in the *negative* framed message condition, when the number of treated patients was lower than the number of those who did not receive the treatment (i.e., in the 100–800 condition), 52% of participants' estimates were *higher* than the exact value compared with 13% and 35% that were lower and accurate, respectively. Denominator neglect accounts for this result, given that the absolute number of patients who received the treatment and died is much lower than the absolute number of patients who did not receive the treatment and died (e.g., 2 and 80, respectively). Focusing on the absolute numbers in the numerators would lead participants to believe that the treatment had a larger effect than it actually did. Instead, in the *positive* framed message condition, there was a tendency to underestimate risk reduction: 48% of participants' estimates were *lower* than the exact value, compared with 9% and 43% that were higher and accurate, respectively. Here, focusing on the absolute numbers in the numerators would lead participants to believe that the treatment had a smaller effect than it actually did.

Instead, when the number of treated patients was higher than the number of patients who did not receive the treatment (i.e., in the 800–100 condition), there was a tendency to underestimate risk reduction in the *negative* framed message condition: 61% of participants' estimates were *lower* than the exact value, compared with 3% and 36% that were higher and accurate, respectively. Instead, in the *positive* framed message condition, 67% of the estimates were *higher* than the exact value, compared with 5% and 28% that were lower and accurate, respectively.

For the 100–800 condition, participants in the negative message framed condition were significantly more likely to overestimate risk reduction, whereas participants in the positive message framed condition were more likely to underestimate risk reduction, $\chi^2(1, n=51)=21.40$, $p=.001$. Instead, for the 800–100 condition, participants in the negative message framed condition were more likely to underestimate risk reduction, whereas participants in the positive message framed condition were more likely to overestimate risk reduction, $\chi^2(1, n=54)=42.55$, $p=.001$.

Next we tested whether icon arrays were more effective in reducing denominator neglect in participants with high graph literacy than in those with low graph literacy (H_1). To this end, we conducted a $2\times2\times4$ ANOVA with icon arrays

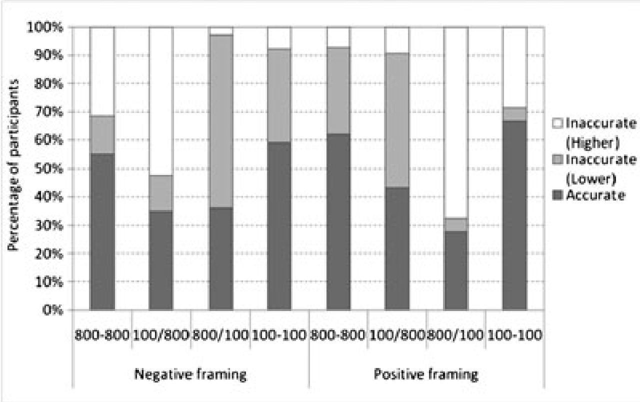

Figure 3. Percentage of participants whose estimates of risk reduction were accurate, lower, or higher than the exact value in the numerical condition only, as a function of the sizes of the denominators and message framing

and graph literacy as between-subjects factors and sizes of denominators as a within-subjects factor on the percentage of participants whose estimates of risk reduction were accurate. Note that in contrast to the first ANOVA reported above, the current analysis included data for participants who did and did not receive icon arrays. Additionally, we excluded message framing as a factor, given the absence of any previous significant effect on denominator neglect. This analysis revealed a main effect of sizes of denominators, $F(3,411)=2.85$, $p=.037$, $\eta_p^2=0.020$, a main effect of icon arrays, $F(1,137)=26.62$, $p=.001$, $\eta_p^2=0.163$, and an interaction between icon arrays and sizes of denominators, $F(3,411)=5.74$, $p=.001$, $\eta_p^2=0.040$. These results indicate that icon arrays helped people to take into account both the overall number of treated and non-treated patients in their estimations of treatment risk reduction. When the sizes of the denominators were different and icon arrays were presented alongside numerical information, the percentage of correct estimates increased from 42% to 73%, and from 34% to 81%, for the 100–800 and 800–100 conditions, respectively ($p<.001$). Instead, when the sizes of denominators were equal, the increase in accuracy when icon arrays were provided was not significant ($p>.1$).

The ANOVA also revealed a significant main effect of graph literacy, $F(1,137)=6.40$, $p=.013$, $\eta_p^2=0.045$, and an interaction between icon arrays and graph literacy, $F(1,137)=4.81$, $p=.030$, $\eta_p^2=0.034$.[2] When icon arrays were not provided, 48% of the participants with low graph literacy provided correct estimates, compared with 64% when icon arrays were provided. For participants with high graph literacy, the percentage of correct estimates instead raised from

[1]An anonymous reviewer suggested the possibility of treating sizes of denominators as a factor with two levels (i.e., same denominators: 800–800 and 100–100 conditions versus different denominators: 800–100 and 100–800 conditions), instead of as a factor with four levels. Two additional ANOVAs conducted following this approach yielded converging results. Specifically, two $2\times2\times2$ ANOVAs with message frame as a between-subjects factor, and sizes of denominators (same versus different) and either sizes of non-treated patients (800 versus 100; first ANOVA) or sizes of treated patients (800 versus 100; second ANOVA) as within-subjects factors revealed only a significant main effect of sizes of denominators ($p<.001$) and no main effect or interaction involving message frame or the sizes of non-treated or of treated patients (Fs<1).

[2]The results of an ANOVA including all factors manipulated yielded converging results. Specifically, a $2\times2\times2\times4$ ANOVA with icon arrays, graph literacy, and message frame as between-subjects factors, and sizes of denominators as a within-subjects factor revealed a main effect of icon arrays ($p=.001$) and of graph literacy ($p=.014$); a marginally significant effect of sizes of denominators ($p=.062$); and interactions between icon arrays and sizes of denominators ($p=.001$) and between icon arrays and graph literacy ($p=.040$). The analysis did not yield a main effect of message frame or any interaction involving this factor.

DOI: 10.1002/bdm

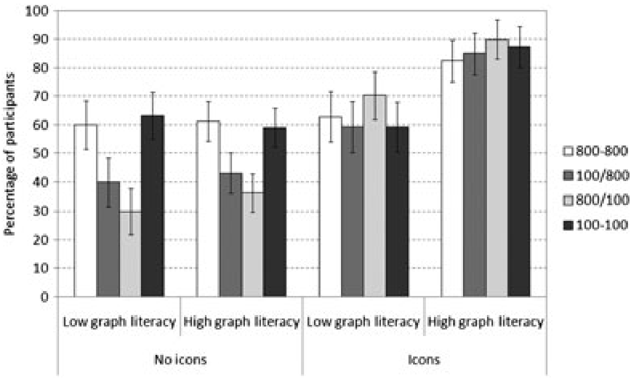

Figure 4. Percentage of participants whose estimates of risk reduction were accurate, as a function of graph literacy, icon arrays, and sizes of the denominators. Error bars represent one standard error

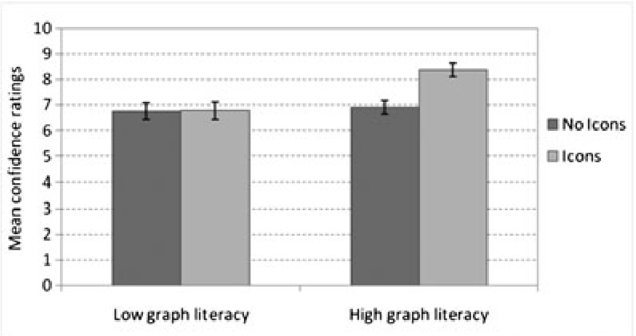

Figure 5. Mean confidence in risk reduction estimates, as a function of graph literacy and icon arrays. Error bars represent one standard error of the mean

51% to 87%. In sum, icon arrays helped all participants (i.e., with high and low graph literacy) to take into account the overall number of treated and non-treated patients, thus reducing denominator neglect. However, in line with H_1, the overall increase in accuracy of risk understanding when icon arrays were provided was significantly larger for participants with high graph literacy than for those with low graph literacy (see Figure 4).[3]

Finally, to test whether differences in graph literacy affected participants' confidence in their estimates of treatment risk reduction (H_2), we conducted a $2 \times 2 \times 4$ ANOVA with icon arrays and graph literacy as between-subjects factors and sizes of denominators as a within-subjects factor on participants' confidence ratings. This analysis revealed a significant main effect of graph literacy, $F(1,164)=7.61$, $p=.006$, $\eta_p^2=0.044$, and icon arrays, $F(1,164)=6.67$, $p=.011$, $\eta_p^2=0.039$, and an interaction between graph literacy and icon arrays, $F(1,164)=5.59$, $p=.019$, $\eta_p^2=0.033$. As Figure 5 shows, the provision of icon arrays alongside the numerical information resulted in a significant increase in confidence for participants with high graph literacy but not for those with low graph literacy, supporting H_2. The mean confidence reported by participants with high graph literacy when only numerical information was provided was 6.9 ($SE=0.3$), whereas it increased to 8.4 ($SE=0.3$) when icon arrays were also presented. Instead, the mean confidence reported by participants with low graph literacy when only numerical information was provided was 6.8 ($SE=0.3$), and 6.8 when icon arrays were provided ($SE=0.3$).

DISCUSSION

A precise understanding of numerical information is essential to accurate judgment and decision making in many contexts. However, numerical concepts can be subject to biases and errors that undermine judgment and decision making. In this paper, we sought to document some of the ways that individual differences in graph literacy can affect the extent to which people benefit from visual aids designed to overcome a common judgment bias (denominator neglect).

Consistent with previous research (e.g., Garcia-Retamero & Galesic, 2009; Garcia-Retamero et al., 2010), we demonstrated that people show denominator neglect when judging the effectiveness of a treatment using information from unequally sized groups of treated and non-treated patients. These results support the idea that ratio concepts are particularly hard to understand (Bonato, Fabbri, Umiltà, & Zorzi, 2007; Ni & Zhou, 2005), and that people tend to behave as if they only compare magnitudes across numerators, thereby neglecting the denominators (Denes-Raj et al., 1995; Reyna, 2004; Reyna & Brainerd, 2008). We also found that icon arrays can help to reduce the effect of denominator neglect. This finding supports the idea that graphical displays are an effective method to reduce judgment biases that can help people to make decisions based on an accurate understanding of risk information. Thus, we support and extend our own and others' previous findings indicating that visual aids often facilitate risk communication in the health domain. Specifically, icon arrays symbolizing individuals who are affected by some risk using circles (Galesic, Garcia-Retamero, & Gigerenzer, 2009; Garcia-Retamero et al., 2010), squares (Zikmund-Fisher et al., 2008), or human figures (Paling, 2003) have been shown to make medical risks easier to interpret. Grids with squares representing visually the specificity or sensitivity of a medical test can also help people to correctly update posttest probabilities (Lloyd & Reyna, 2001). Visual aids can also improve understanding of risks associated with different medical treatments, screenings, and lifestyles (Ancker et al., 2006; Galesic et al., 2009; Garcia-Retamero & Galesic, 2010b; Garcia-Retamero, Galesic, & Gigerenzer, 2011; Lipkus, 2007); promote consideration of beneficial treatments that have side effects (Waters, Weinstein, Colditz, & Emmons, 2007); and eliminate

[3]The dichotomization of a continuous measure by median split can have negative consequences such as the loss of statistical power. Therefore, we also performed logistic regressions using the full range of graph literacy scores as a predictor and accuracy as a dependent variable, for each of the four sizes of denominators conditions. In all cases, a test of the model versus a model with intercept only was statistically significant when icon arrays were provided ($\chi^2=3.91–12.53$, $df=1$, $p=.001–.048$, odds ratio=1.26–1.59) but not when only numerical information was provided ($p>.20$ for all tests). These results are in line with those obtained in the ANOVA reported, where graph literacy was dichotomized via a median split.

DOI: 10.1002/bdm

errors induced by anecdotal narratives (Fagerlin, Wang, & Ubel, 2005).

Additionally, we found that the increase in the percentage of participants providing accurate estimates when icon arrays were provided was larger when the sizes of the denominators were different, than when they were equal. Considering the information in the denominators is crucial to calculate risk reduction accurately when the sizes of denominators differ (Garcia-Retamero & Galesic, 2009; Garcia-Retamero et al., 2010). Thus, interventions such as icon arrays that help people to take denominators into account can be particularly effective in such cases.

Our results extended previous research in at least two other notable ways. First, we demonstrated that individual differences in graph literacy can moderate the effectiveness of icon arrays in increasing accuracy of risk understanding. Icon arrays helped to reduce denominator neglect both for participants with high and low graph literacy. However, the increase in risk comprehension was larger for highly graph-literate participants than for less graph-literate participants. Second, we established that individual differences in graph literacy can affect people's confidence in their estimates of risk reduction. The provision of icon arrays resulted in a significant increase in highly graph-literate participants' self-reported confidence in their own estimates, whereas less graph-literate participants' confidence was not reliably affected. Finally, our results also provided evidence inconsistent with the suggestion that an attentional bias might exacerbate the effect of denominator neglect when information is framed negatively, given that the percentage of accurate responses did not vary as a function of the type of framing provided. These results suggest that denominator neglect can hold consistently, independent of whether risk information is presented in positive or negative terms.

Why do visual aids improve accuracy in risk understanding? Potential explanatory mechanisms and open questions for future research

Theoretical frameworks such as the Fuzzy Trace Theory (e.g., Brainerd & Reyna, 1990; Reyna & Brainerd, 1995; Reyna, et al., 2009) have been used to provide an explanation of the potential mechanisms underlying the effects of icon arrays on denominator neglect. According to previous data, the problems that people face to understand ratio concepts stem from the fact that the references of classes overlap, which leads to class-inclusion errors (Brainerd & Reyna, 1990; Reyna, 1991). Thus, denominator neglect would be produced by people's tendency to focus on the target classes in numerators, thereby neglecting classes in denominators. Previous research suggests that manipulations that contribute to disentangle classes (e.g., icon arrays) can help to reduce biases such as denominator neglect as a function of shifting processing from more verbatim based to gist-based representation of set structures (Brainerd & Reyna, 1990; Reyna et al., 2009).

A related explanation of the power of icon arrays to reduce denominator neglect was put forward by Stone et al. (2003; see also Ancker et al., 2006). Stone and colleagues suggested that Yamagishi's (1997) findings illustrating denominator neglect can be explained in terms of the saliency of foreground information (e.g., number of people harmed, or subset) versus background information (e.g., the number of people at risk, or superordinate set). Thus, according to Stone et al., the numerical presentation of information can lead to a focus on foreground information, whereas graphical formats displaying both foreground and background information—pie charts and stacked bar graphs—contribute to bringing people's attention to the background too.

Our findings are compatible with the hypotheses that (i) icon arrays contribute to disentangle classes and (ii) icon arrays bring people's attention to background information. As a consequence, these kinds of displays are likely to help people to overcome denominator neglect in a variety of common situations. The mechanisms outlined by Brainerd and Reyna (1990; see also Reyna et al., 2009) and Stone et al. (2003) anticipate the effectiveness of icon arrays in reducing denominator neglect observed in our study. However, the interactions between icon arrays and graph literacy obtained here (for both accuracy and confidence) suggest that (i) the power of icon arrays to increase the accuracy of risk reduction estimates is larger for highly graph-literate individuals than for less graph-literate ones and that (ii) the subjective perceptions of individuals with low graph literacy (i.e., confidence ratings) are not necessarily influenced by the presence of icon arrays. These findings are compatible with our hypothesis that a certain level of graph literacy can be necessary in order to associate the visual patterns contained in icon arrays with meaningful interpretations of the data represented (i.e., risk reduction information).

Future work should directly aim to trace attentional and cognitive processes underlying the effect of icon arrays in individuals with different levels of graph literacy. The use of process tracing methodologies such as eye-tracking or verbal protocol analysis would provide a more nuanced understanding of the time course and operations involved for participants of varying skill levels. For example, the analysis of eye movement data would assess the proportional fixation times on the circles in the icon arrays representing number of patients who died (foreground information) versus on circles representing number of people at risk (background information). These data might reveal differences in the saliency of foreground versus background information in icon arrays for different viewers. Additionally, process tracing would enable the testing and refining of higher fidelity cognitive process theories that explain or, more importantly, predict how various kinds of displays are processed by different viewers. This would expand previous research that has documented graph comprehension processes (e.g., encoding of the visual pattern, translation of visual features into conceptual relations) mainly in homogeneous viewers, focusing on displays that include features such as axes or scales (e.g., line plots or bar charts; Carpenter & Shah, 1998; Lohse, 1993; Pinker, 1990).

From a translational or applied standpoint, process tracing is an essential step in efforts to facilitate the development of training methods for individuals with low graph literacy. These methods could be based in part on the processes that

DOI: 10.1002/bdm

highly graph-literate individuals follow to understand these kinds of displays. That is, process-tracing studies allow for "reverse engineering" of superior performance by revealing encoding or search strategies of successful individuals that may confer benefits to those participants who do not yet use such strategies (Cokely, Kelley, & Gilchrist, 2006; Cokely & Kelley, 2009). Moreover, an understanding of the encoding and search processes of low performing individuals may also provide clues for the design of environments that facilitate more appropriate search and representation. Nevertheless, many open questions remain concerning when and for whom simple differences in encoding or representational strategies (e.g., gist-based representation) would be sufficient for improved performance. Ongoing research is investigating these issues (Okan, Galesic, & Garcia-Retamero, 2010).

Theoretical and practical implications

Taken together, our findings show that visual aids do not necessarily facilitate risk comprehension to the same extent for everyone. Our results emphasize the importance of considering the fit between (i) persons; (ii) cognitive processes; and (iii) task environments when designing interventions such as visual aids. Individual differences in graph literacy moderate the effect of such visual aids, affecting the accuracy of risk judgments. Similarly, Parker and Fischhoff (2005) identified a set of tasks that capture four basic skills required by competent decision makers. One of these skills (*belief assessment*) refers to people's ability to judge the probability of occurrence of events. Our data indicate that the accurateness of this kind of judgment may be affected by variations in graph literacy, suggesting that graph literacy is usefully characterized as a cognitive skill that influences competent decision making.

A growing body of research has documented a variety of individual differences that influence decision making performance. These include domain general decision making skills such as those identified by Parker and Fischhoff (see also Bruine de Bruin et al., 2007; Finucane, Mertz, Slovic, & Schmidt, 2005) or skills such as numeracy (Peters & Levin, 2008; Peters et al., 2006). Other individual differences that can have an effect on the quality of risky judgment and decision making include decision making styles (Baron, 2000; Campitelli & Labollita, 2010; Frederick, 2005; Shiloh, Salton, & Sharabi, 2002), specific expertise (Ericsson, Prietula, & Cokely, 2007; Garcia-Retamero & Dhami, 2009; Shanteau, 1992), and domain general cognitive abilities (Cokely & Kelley, 2009; Del Missier, Mäntylä, & Bruine de Bruin, 2010, in press; Stanovich & West, 2000, 2008). Research indicates that general decision making skills have significant relations among them and with other measures of cognitive abilities and styles (Bruine de Bruin et al., 2007; Del Missier et al., 2010; Parker & Fischhoff, 2005). Future research should aim to achieve a more precise specification of the relations between graph literacy, the set of individual differences outlined above, and the cognitive processes that mediate differences.

The current findings also provide additional support for the predictive validity of the scale developed by Galesic and Garcia-Retamero (2011; see also Garcia-Retamero & Galesic, 2010b). The evidence obtained so far suggests that this scale can be useful to predict performance in tasks involving not only icon arrays but also other kinds of graphs. Future work should examine the extent to which graph literacy moderates performance in tasks involving displays such as pie charts or line plots.

Future research should aim to enhance (i) the generalizability and (ii) the ecological validity of the present study. Concerning the first point, it should be noted that the instrument and the materials used focused on the medical domain. Thus, the extent to which graph literacy can influence risky decision making performance in other important domains, such as finance or politics, is not yet well documented or understood. However, ongoing studies do seem to suggest that graph literacy will have some predictive validity across diverse domains. Concerning ecological validity, the fact that our experiments were not conducted in a clinical setting prompts us to suggest some caution regarding immediate prescriptive applications of our findings for medical practice. Research has shown that the effect of manipulations observed in the laboratory, such as framing of information in the context of risk communication, may not be generalizable to clinical practice (Edwards et al., 2001). Thus, future research should aim to provide more converging evidence on the effect of graph literacy in the ecology of interest (e.g., in the clinic).

Conclusion

In the present article, we have demonstrated that individual differences in graph literacy can moderate the magnitude of the effect of visual risk communication interventions (i.e., icon arrays). This finding is relevant to modern societies, where graphical displays are increasingly being used and recommended for the communication of risks to the public (Ancker et al., 2006; Fuller et al., 2002; Lipkus, 2007). However, these graphs are rarely designed on the basis of a systematic set of principles of good graph construction. As a consequence, data represented in graphs are frequently misinterpreted by viewers (Beattie & Jones, 2002; Cooper, Schriger, Wallace, Mikulich, & Wilkes, 2003). This can lead to substantial alterations in viewers' preferences and decisions (Arunachalam, Pei, & Steinbart, 2002). Principles of good graph design have been put forward in some cases (Cleveland, 1994; Cleveland & McGill, 1984; Kosslyn, 2006; Tufte, 2001), but a consensus does not always exist regarding the adequacy or effective implementation of these principles. For instance, Kosslyn (2006) argues that it is adequate to use bar charts in which part of the *y*-axis has been removed, provided that marks are used to indicate discontinuities. In contrast, other authors suggest that the scale in the *y*-axis of bar charts should start at 0, in order for the proportion between the lengths of the bars to reflect the proportion between the quantitative data (e.g., Cleveland, 1994).

There is a need for a unified and usable set of standards for guiding graph designers' work. We suspect the timing

DOI: 10.1002/bdm

is right to work toward refining the available guidelines that have been developed across several fields (e.g., cognitive psychology, mathematics, human factors). Resulting standards could play an important role in inoculating professionals, policy makers, and the general public against potentially distorted and misleading communications. As highlighted by the current data, graph design standards should allow for custom-tailored designs that are sensitive to the various needs and abilities of diverse individuals. Additionally, the scope of the effect of graph literacy on a range of different judgment and decision making tasks should be investigated, along with relations to other decision making competences. To an important extent, such goals hinge on our ability to achieve a more unified understanding of the theoretical and practical relevance of the construct of graph literacy, including documentation of both construct and predictive validity of the instruments used to measure it. This work holds the promise of important theoretical and translational benefits such as improved decision making in medicine, business, and potentially many other domains involving the communication of risk.

REFERENCES

Ancker, J. S., Senathirajah, Y., Kukafka, R., & Starren, J. B. (2006). Design features of graphs in health risk communication: A systematic review. *Journal of the American Medical Informatics Association*, *13*, 608–618.

Arunachalam, V., Pei, B. K., & Steinbart, P. J. (2002). Impression management with graphs: Effects on choices. *Journal of Information System*, *16*, 183–202.

Banks, S. M., Salovey, P., Greener, S., Rothman, A. J., Moyer, A., Beauvais, J., & Epel, E. (1995). The effects of message framing on mammography utilization. *Health Psychology*, *14*, 178–184.

Baron, J. (2000). *Thinking and deciding* (3rd ed.).Cambridge, England: Cambridge University Press.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323–370.

Beattie, V., & Jones, M. J. (2002). Measurement distortion of graphs in corporate reports: An experimental study. *Accounting, Auditing & Accountability Journal*, *15*, 546–564.

Bonato, M., Fabbri, S., Umiltà, C., & Zorzi, M. (2007). The mental representation of numerical fractions: Real or integer? *Journal of Experimental Psychology. Human Perception and Performance*, *33*, 1410–1419.

Bonner, C., & Newell, B. R. (2008). How to make a risk seem riskier: The ratio bias versus construal level theory. *Judgment and Decision Making*, *3*, 411–416.

Brainerd, C. J., & Reyna, V. F. (1990). Inclusion illusions: Fuzzy-trace theory and perceptual salience effects in cognitive development. *Developmental Review*, *10*, 365–403.

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision making competence. *Journal of Personality and Social Psychology*, *92*, 938–956.

Campitelli, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making*, *5*, 182–192.

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology. Applied*, *4*, 75–100.

Cleary, P. D., & Angel, R. (1984). The analysis of relationships involving dichotomous dependent-variables. *Journal of Health and Social Behavior*, *25*, 334–348.

Cleveland, W. S. (1994). *The elements of graphing data*. Summit, NJ: Hobart Press.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*, 531–554.

Cokely, E. T., Galesic, M., Schulz, E., & Garcia-Retamero, R. (2011). *The Berlin advanced numeracy test for highly educated samples (ANT-E): A valid three minute test of statistical and risk literacy*. Manuscript submitted for publication.

Cokely, E. T., Ghazal, S., Garcia-Retamero, R., Galesic, M., & Schulz, E. (in press). Cross-Cultural comparisons of statistical numeracy in college educated samples. In R. Garcia-Retamero, & M. Galesic (Eds.), *Risk communication and medical decision making: A cross-cultural perspective*. New York: Springer.

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.

Cokely, E. T., Kelley, C. M., & Gilchrist, A. H. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin & Review*, *13*, 991–997.

Cooper, R. J., Schriger, D. L., Wallace, R. C., Mikulich, V. J., & Wilkes, M. S. (2003). The quantity and quality of scientific graphs in pharmaceutical advertisements. *Journal of General Internal Medicine*, *18*, 294–297.

Del Missier, F., Mäntylä, T., & deBruine, W. B. (2010). Executive functions in decision making: An individual differences approach. *Thinking and Reasoning*, *16*, 69–97.

Del Missier, F., Mäntylä, T.,& de Bruin, W. B. (in press). Decision making competence, executive functioning, and general cognitive abilities. *Journal of Behavioral Decision Making*. doi:10.1002/bdm.731 (this issue).

Denes-Raj, V., Epstein, S., & Cole, J. (1995). The generality of the ratio-bias phenomenon. *Personality and Social Psychology Bulletin*, *21*, 1083–1092.

Edwards, A., Elwyn, G., Covey, J., Matthews, E., & Pill, R. (2001). Presenting risk information a review of the effects of framing and other manipulations on patient outcomes. *Journal of Health Communication*, *6*, 61–82.

Edwards, A., Elwyn, G., & Mulley, A. (2002). Explaining risks: Turning numerical data into meaningful pictures. *British Medical Journal*, *324*, 827–830.

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, *85*, 114–121.

Fagerlin, A., Wang, C., & Ubel, P. (2005). Reducing the influence of anecdotal reasoning on people's health care decisions: Is a picture worth a thousand statistics? *Medical Decision Making*, *25*, 398–405.

Finucane, M. L., Mertz, C. K., Slovic, P., Schmidt, E. S. (2005). Task complexity and older adults' decision making competence. *Psychology and Aging*, *20*, 71–84.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 24–42

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*, 124–158.

Fuller, R., Dudley, N., & Blacktop, J. (2002). How informed is consent? Understanding of pictorial and verbal probability information by medical inpatients. *Postgraduate Medical Journal*, *78*, 543–544.

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.

Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making*, *31*, 444 –457.

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., & Cokely, E. T. (in press). Effective communication of risks to young adults: Using message framing and visual aids to increase condom use and STD screening. *Journal of Experimental Psychology: Applied*. doi:10.1037/a0023677

Garcia-Retamero, R., & Dhami, M. K. (2009). Take-the-best in expert-novice decision strategies for residential burglary. *Psychonomic Bulletin and Review*, *16*, 163–169.

Garcia-Retamero, R., & Dhami, M. K. (2011). Pictures speak louder than numbers: On communicating medical risks to immigrants with limited non-native language proficiency. *Health Expectations*, *14*, 46–57.

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.

Garcia-Retamero, R., & Galesic, M. (2010a).How to reduce the effect of framing on messages about health. *Journal of General Internal Medicine*, *25*, 1323–1329.

Garcia-Retamero, R., & Galesic, M. (2010b). Who profits from visual aids: Overcoming challenges in people's understanding of risks. *Social Science & Medicine*, *70*, 1019–1025.

Garcia-Retamero, R., & Galesic, M. (2011). Using plausible group sizes to communicate information about medical risks. *Patient Education and Counseling*, *84*, 245–250.

Garcia-Retamero, R., Galesic, M., & Gigerenzer, G. (2010). Do icon arrays help reduce denominator neglect? *Medical Decision Making*, *30*, 672–684.

Garcia-Retamero, R., Galesic, M., Gigerenzer, G. (2011). Enhancing understanding and recall of quantitative information about medical risks: A cross-cultural comparison between Germany and Spain. *Spanish Journal of Psychology*, *14*, 218–226.

Gattis, M. (2002). Structure mapping in spatial reasoning. *Cognitive Development*, *17*, 1157–1183.

Gattis, M. (2004). Mapping relational structure in spatial reasoning. *Cognitive Science*, *28*, 589–610.

Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*, 231–239.

Gerend, M. A., & Shepherd, J. E. (2007). Using message framing to promote acceptance of the human papillomavirus vaccine. *Health Psychology*, *26*, 745–752.

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2007). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, *8*, 53–96.

Grossarth-Maticek R., & Ziegler R. (2008). 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. *European Journal of Medical Research*, *13*, 107–120.

Hansen, C., & Hansen, R. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology*, *54*, 917–924.

Haward, M. F., Murphy, R. O., & Lorenz, J. M. (2008). Message framing and perinatal decisions. *Pediatrics*, *122*, 109–118.

Kosslyn, S. M. (2006). *Graph design for the eye and mind*. New York, NY: Oxford University Press.

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, *76*, 149–188.

Lichtenberg P., Levinson D., Sharshevsky Y., Feldman D., & Lachman, M. (2008). Clinical case management of revolving door patients: A semi-randomized study. *Acta Psychiatrica Scandinavica*, *117*, 449–454.

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.

Lloyd, F. J., & Reyna, V. F. (2001). A Web exercise in evidence-based medicine using cognitive theory. *Journal of General Internal Medicine*, *16*, 94–99.

Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human Computer Interaction*, *8*, 353–388.

Lunney, G. H. (1970). Using analysis of variance with a dichotomous dependent variable—An empirical study. *Journal of Educational Measurement*, *7*, 263–269.

Maichle, U. (1994). Cognitive processes in understanding line graphs. In W. Schnotz & R. W. Kulhavy(Eds.),*Comprehension of graphics* (pp. 207–227). Amsterdam: North-Holland Elsevier Science

Marteau, T. M. (1989). Framing of information: Its influence upon decisions of doctors and patients. *British Journal of Social Psychology*, *28*, 89–94.

Mazur, D. J., & Hickam, D. H. (1993). Patients' and physicians' interpretations of graphic data displays. *Medical Decision Making*, *13*, 59–63.

McNeil, B. J., Pauker, S. G., Sox, H. C., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *The New England Journal of Medicine*, *306*, 1259–1262.

Ni, Y., & Zhou, Y. D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, *40*, 27–52.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*, 381–396.

Okan, Y., Galesic, M., & Garcia-Retamero, R. (2010). *Graph comprehension in medical contexts: An eye-tracking study*. Poster presented at the 13th Biennial European Conference of the Society for Medical Decision Making, Hall in Tyrol.

Paling, J. (2003). Strategies to help patients understand risks. *British Medical Journal*, *327*, 745–748.

Parker, A. M., & Fischhoff, B. (2005). Decision making competence: External validation through an individual-differences approach. *Journal of Behavioral Decision Making*, *18*, 1–27.

Peters, E. & Levin, I. P. (2008). Dissecting the risky-choice framing effect: Numeracy as an individual-difference factor in weighting risky and riskless options. *Judgment and Decision Making*, *3*, 435–448.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, *17*, 407–413.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73–126). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*, 380–391.

Reyna, V. F. (1991). Class inclusion, the conjunction fallacy, and other cognitive illusions. *Developmental Review*, *11*, 317–336.

Reyna, V. F. (2004). How people make decisions that involve risk. *Current Directions in Psychological Science*, *13*, 60–66.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*, 1–75.

Reyna, V. F., & Brainerd, C. J. (2008).Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*, 89–107.

Reyna, V. F., Nelson, W. L., Han P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, *135*, 943–973.

Rivers, S. E., Salovey, P., Pizarro, D. A., Pizarro, J., & Schneider, T. R. (2005). Message framing and pap test utilization among women attending a community health clinic. *Journal of Health Psychology*, *10*, 65–77.

Roth, W. M., & Bowen, G. M. (2003). When are graphs worth ten thousand words? An expert–expert study. *Cognition and Instruction*, *21*, 429–473.
Rothman, A. J., Martino, S. C., Bedell, B. T., Detweiler, J. B., & Salovey, P. (1999). The systematic influence of gain-and loss-framed messages on interest in and use of different types of health behavior. *Personality and Social Psychology Bulletin*, *25*, 1355–1369.
Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behavior: The role of message framing. *Psychological Bulletin*, *121*, 3–19.
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.
Shah, P. (2001). Graph comprehension: The role of format, content, and individual differences. In M. Anderson, B. Meyer, & P. Olivier (Eds.), *Diagrammatic representation and reasoning*. (pp 173–185). Berlin: Springer-Verlag.
Shah, P., & Freedman, E. G. (2011). Bar and line graph comprehension: An interaction of top-down and bottom-up processes. *Topics in Cognitive Science*, *3*, 560–578.
Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review*, *14*, 47–69.
Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, *81*, 75–86.
Shiloh, S., Salton, E., & Sharabi, D. (2002). Individual differences in rational and intuitive thinking styles as predictors of heuristic responses and framing effects. *Personality and Individual Differences*, *32*, 415–429.
Smith, N. K., Cacioppo, J. T., Larsen, J. T., & Chartrand, T. L. (2003). May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsycologia*, *41*, 171–183.
Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–726.
Stanovich, K. E., & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology*, *94*, 672–695.
Stone, E. R., Sieck, W. R., Bull, B. E., Yates, J. F., Parks, S. C., & Rush, C. J. (2003). Foreground: Background salience: Explaining the effects of graphical displays on risk avoidance. *Organizational Behavior and Human Decision Processes*, *90*, 19–36.
Toll, B. A., Martino, S., Latimer, A., Salovey, P., O'Malley, S., Calin-Menter, S., Hopkins J., Wu, R., Celestino, P., & Cummings, K. M. (2010). Randomized trial: Quitline specialist training in gain-framed vs standard-care messages for smoking cessation. *Journal of the National Cancer Institute*, *102*, 96–16.
Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Cheshire, CT: Graphics Press.
Tversky, B. (2001). Spatial schemas in depictions. In M. Gattis (Ed.), *Spatial schemas and abstract thought* (pp. 79–112). Cambridge, MA: MIT Press.
Waters, E. A., Weinstein, N. D., Colditz, G. A., & Emmons, K. M. (2007). Reducing aversion to side effects in preventive medical treatment decisions. *Journal of Experimental Psychology. Applied*, *13*, 11–21.
Wilson, D. K., Kaplan, R. M., & Schneiderman, L. J. (1987). Framing of decisions and selections of alternatives in health care. *Social Behaviour*, *2*, 51–59.
Wright, A. J., Whitwell, S. C., Takeichi, C., Hankins, M., & Marteau, T. M. (2009). The impact of numeracy on reactions to different graphic risk presentation formats: An experimental analogue study. *British Journal of Health Psychology*, *14*, 107–125.
Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: Implications for risk communication. *Applied Cognitive Psychology*, *11*, 495–506.
Zikmund-Fisher, B., Ubel, P., Smith, D., Derry, H., McClure, J., Stark, A., Pitsch, R., & Fagerlin, A. (2008). Communicating side effect risks in a tamoxifen prophylaxis decision aid: The debiasing influence of pictographs. *Patient Education and Counseling*, *73*, 209–214.

Authors' biographies:

Yasmina Okan is a Ph.D. Candidate at the University of Granada in Spain.

Rocio Garcia-Retamero is a Professor at the University of Granada in Spain. She is also an Associated Scientist with the Center for Adaptive Behavior and Cognition, at the Max Planck Institute for Human Development, in Berlin, Germany.

Edward T. Cokely is an Assistant Professor of Psychology and Director of the Decision Science and Decision Engineering Laboratory at the Michigan Technological University. He is also an Associated Scientist with the Center for Adaptive Behavior and Cognition, at the Max Planck Institute for Human Development, in Berlin, Germany.

Antonio Maldonado is a Full Professor at the University of Granada in Spain and Director of the research group Learning, Emotion, and Decision.

Authors' addresses:

Yasmina Okan, Rocio Garcia-Retamero and **Antonio Maldonado**, Department of Experimental Psychology, University of Granada, Granada, Spain.

Edward T. Cokely, Department of Cognitive and Learning Sciences, Michigan Technological University, Houghton, MI, USA.

DOI: 10.1002/bdm