Designing Visual Aids That Promote Risk Literacy: A Systematic Review of Health Research and Evidence-Based Design Heuristics

Rocio Garcia-Retamero, University of Granada, Granada, Spain, and **Edward T. Cokely**, University of Oklahoma, Norman

Background: Effective risk communication is essential for informed decision making. Unfortunately, many people struggle to understand typical risk communications because they lack essential decision-making skills.

Objective: The aim of this study was to review the literature on the effect of numeracy on risk literacy, decision making, and health outcomes, and to evaluate the benefits of visual aids in risk communication.

Method: We present a conceptual framework describing the influence of numeracy on risk literacy, decision making, and health outcomes, followed by a systematic review of the benefits of visual aids in risk communication for people with different levels of numeracy and graph literacy. The systematic review covers scientific research published between January 1995 and April 2016, drawn from the following databases: Web of Science, PubMed, PsycINFO, ERIC, Medline, and Google Scholar. Inclusion criteria were investigation of the effect of numeracy and/or graph literacy, and investigation of the effect of visual aids or comparison of their effect with that of numerical information. Thirty-six publications met the criteria, providing data on 27,885 diverse participants from 60 countries.

Results: Transparent visual aids robustly improved risk understanding in diverse individuals by encouraging thorough deliberation, enhancing cognitive self-assessment, and reducing conceptual biases in memory. Improvements in risk understanding consistently produced beneficial changes in attitudes, behavioral intentions, trust, and healthy behaviors. Visual aids were found to be particularly beneficial for vulnerable and less skilled individuals.

Conclusion: Well-designed visual aids tend to be highly effective tools for improving informed decision making among diverse decision makers. We identify five categories of practical, evidence-based guidelines for heuristic evaluation and design of effective visual aids.

Keywords: data visualization, graph literacy, graphic usability, heuristic evaluation, numeracy, risk communication, risk literacy, shared decision making, visual aids

Address correspondence to Rocio Garcia-Retamero, PhD, Facultad de Psicologia, Universidad de Granada, Campus Universitario de Cartuja s/n, 18071 Granada, Spain; e-mail: rretamer@ugr.es.

HUMAN FACTORS
Vol. XX, No. X, Month XXXX, pp. 1–46
DOI: 10.1177/0018720817690634

INTRODUCTION

Informed decision making depends on one's ability to accurately evaluate and understand information about risk—that is, *risk literacy* (Ancker & Kaufman, 2007; Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Peters, 2012; Reyna, Nelson, Han, & Dieckmann, 2009). For instance, physicians need to interpret and communicate information about the benefits and risks of different medical treatments, screenings, and lifestyle choices (Anderson & Schulkin, 2014b), and patients need to understand and use this information to consent on their own behalf to medical treatment and to adhere to medical advice (Zikmund-Fisher, Mayman, & Fagerlin, 2014). Unfortunately many people, including highly trained health care professionals and motivated patients, do not have the skills necessary for accurate, unaided evaluation of risks (Cokely et al., 2012; Garcia-Retamero & Galesic, 2013).

To illustrate, consider a recent systematic review by Anderson and Schulkin (2014b), who investigated physicians' understanding and use of health-relevant numerical information. In the studies reviewed, numeracy—that is, practical probabilistic reasoning skill—was assessed using an 11-question test developed by Schwartz, Woloshin, Black, and Welch (1997) and Lipkus, Samsa, and Rimer (2001). This efficient test has been used in over 100 studies on various topics. The test specifically assesses practical knowledge of basic probabilistic concepts and mathematical operations, including comparing risk magnitudes, converting percentages to proportions, converting proportions to percentages, converting probabilities to proportions, and computing probabilities.

Examples of items from the test are "If Person A's risk of getting a disease is 1% in ten years, and Person B's risk is double that of A's, what is

B's risk?" "Which of the following numbers represents the biggest risk of getting a disease? 1 in 10, 1 in 100, or 1 in 1,000?" and "If the chance of getting a disease is 20 out of 100, this would be the same as having a __% chance of getting the disease." Surprisingly, Anderson and Schulkin (2014b) found that 25% to 47% of the sampled physicians could not correctly answer all of these basic numeracy questions. Other researchers have since documented similar findings in physicians (Anderson, Obrecht, Chapman, Driscoll, & Schulkin, 2011; Estrada, Barnes, Collins, & Byrd, 1999; Garcia-Retamero, Wicki, Cokely, & Hanson, 2014), medical students (Hanoch, Miron-Shatz, Cole, Himmelstein, & Federman, 2010), and nurses (Lopez et al., 2016).

Research drawing on general community samples further documents the common struggles people have with basic numerical concepts (Ancker & Kaufman, 2007; Cokely, Ghazal, & Garcia-Retamero, 2014; Fagerlin, Ubel, Smith, & Zikmund-Fisher, 2007; Peters, 2012; Reyna et al., 2009). For instance, Lipkus et al. (2001) conducted a series of studies on community samples of U.S. veterans, who answered the 11 numeracy questions described earlier. Results showed that even among this relatively educated sample, many individuals were statistically innumerate: 20% of the participants incorrectly answered questions dealing with risk magnitude, and 30% were unable to correctly infer that if 20 in 100 people get a disease, the chance of suffering the disease would be 20%. A more recent study conducted by Galesic and Garcia-Retamero (2010) documented similar findings using large probabilistic (i.e., representative) national samples of the populations of the United States and Germany. These results suggest that more than 25% of all adults living in these countries are likely to have profound numeracy processing deficits. For instance, 28% and 25% of the residents in Germany and the United States, respectively, were unable to accurately infer that a risk that affects 1 in 10 people is larger than one that affects 1 in 100 or 1 in 1,000 people. Roughly 17% of residents also incorrectly answered the easiest question (i.e., "If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 1,000?").

In addition, Cokely, Garcia-Retamero, and colleagues have investigated numeracy and risk literacy in more than 10,000 people from 60 countries (e.g., China, Japan, India, Pakistan, Sweden, Spain, England, and the United States), including broad samples of professionals, probabilistic national samples, young adults, patients, doctors, police officers, managers, college students, and Web panels (Cokely et al., 2012; Garcia-Retamero, Cokely, & Ghazal, 2016; Garcia-Retamero et al., 2014; Ghazal, Cokely, & Garcia-Retamero, 2014). Across all studies, results converge, indicating that many people do not possess basic, practical mathematical skills that are often essential for independent risk evaluation and general skilled decision making (for a review, see Cokely et al., in press).

Evidence documenting considerable rates of innumeracy in industrialized countries also comes from several extensive longitudinal studies, such as the Programme for International Student Assessment (PISA) and the National Assessment of Adult Literacy (NAAL) (Kutner, Greenburg, Jin, & Paulsen, 2006; Organisation for Economic Co-operation and Development, 2012). For instance, results from the NAAL indicated that 22% of American adults fall within the lowest level of quantitative literacy (i.e., below basic level). This percentage is even larger (29%) among the group of uninsured adults (Peters, Meilleur, & Tompkins, 2014). This suggests that about 1 in 3 newly insured individuals are unlikely to correctly interpret familiar numerical expressions of probability or perform simple quantitative operations (e.g., addition; Peters et al., 2014).

Perhaps even more problematically, Rodríguez et al. (2013) conducted a study on U.S. patients' understanding of numerical information, finding that on average, patients were older, less educated, and less numerate than the general population: More than 40% of the patients in this study incorrectly answered most of the items developed by Schwartz et al. (1997) and Lipkus et al. (2001) (for similar results, see Abdel-Kader et al., 2010; Osborn, Cavanaugh, Wallston, & Rothman, 2010; Sharit et al., 2014; Zikmund-Fisher, Mayman, et al., 2014).

Although many people may be at a disadvantage when it comes to some risk communication practices, a growing body of research suggests that simple visualization technologies—that is, well-designed visual aids—can often dramatically improve risk communication, comprehension, and skilled decision making among diverse users (Garcia-Retamero & Cokely, 2013, 2014b). To thoroughly examine this issue and extract essential lessons, we conducted a systematic review and an analysis of the available literature concentrated in health and medical decision making. Our review begins with a basic introduction to findings on the robust link between skills and quality outcomes. We next present an evidence-based conceptual framework describing psychological, social, and technological factors that shape the influence of numeracy on risk literacy, decision making, and health outcomes. Following the recommendations of the Cochrane handbook for systematic reviews (Higgins & Green, 2008), we then present a systematic review of research investigating the influence of skills on the benefits of visual aids.

In the systematic review we consider (a) the consistency of evidence on the benefits of visual aids with emphasis on risk understanding, decision making, and healthy behavior; (b) an evaluation of the benefits of visual aids across types of visual aids, types of risks, and levels of interactivity; (c) a framework explaining the cognitive and behavioral mechanisms that give rise to superior decision making; (d) a review of subjective perceptions and opinions about the value and utility of various types of visual aids; and (e) a discussion of effective visual aid design, including a five-category group of evidence-based design heuristics. We close with a discussion of open questions, emerging opportunities, and policy implications.

NUMERACY, DECISION MAKING, AND HEALTH: A CONCEPTUAL FRAMEWORK

A robust link between numeracy and health outcomes has been established in many health-related studies on topics such as disease management, quality of life, risk of hospitalization, and prevalence of comorbidity. Low levels of numeracy predict patients' difficulty following a complex dosing regimen (Estrada, Martin-Hryniewicz, Peek, Collins, & Byrd, 2004; Waldrop-Valverde, Jones, Gould, Kumar, & Ownby, 2010), their total body mass index (Huizinga, Beech, Cavanaugh, Elasy, & Rothman, 2008), and their activity limitations as well as struggles with emotional functioning (Apter et al., 2009). Low levels of numeracy also predict patients' utilization of emergency department services (Apter et al., 2006; Ginde, Clark, Goldstein, & Camargo, 2008) and longer delays when seeking medical attention, which can dramatically increase patients' risk for death and major disability (Petrova et al., in press).

Compared with patients with more developed numerical skills, less numerate patients are also at greater risk for comorbid conditions, including myocardial infarction, chronic obstructive pulmonary disease, peptic ulcer disease, liver disease, diabetes, and HIV/AIDS (Garcia-Retamero, Andrade, Sharit, & Ruiz, 2015; Petrova et al., in press). On average, the relative risk that patients with low numeracy suffer one of these diseases is roughly 40% greater than that of patients with high numeracy. Moreover, patients with low numeracy take 20% more prescribed medication than those with higher numerical skills (Garcia-Retamero, Andrade, et al., 2015). These findings hold statistically after controlling for the effect of key demographics (e.g., age, education, ethnicity, and household income) and body mass index, suggesting that numeracy has a unique effect on health outcomes. Theoretically, numeracy may affect the prevalence of comorbidity and medication intake via several mediating pathways (see the conceptual framework in Figure 1).

Generally, the single most influential factor that mediates the relations between numeracy and prevalence of comorbidity is the accuracy and adequacy of patients' risk understanding. Patients with low numeracy often have less accurate perceptions of the benefits and risks of medical treatments and interventions. Compared with patients with high numeracy, less numerate patients overestimate their personal risk of suffering several diseases (Davids, Schapira,

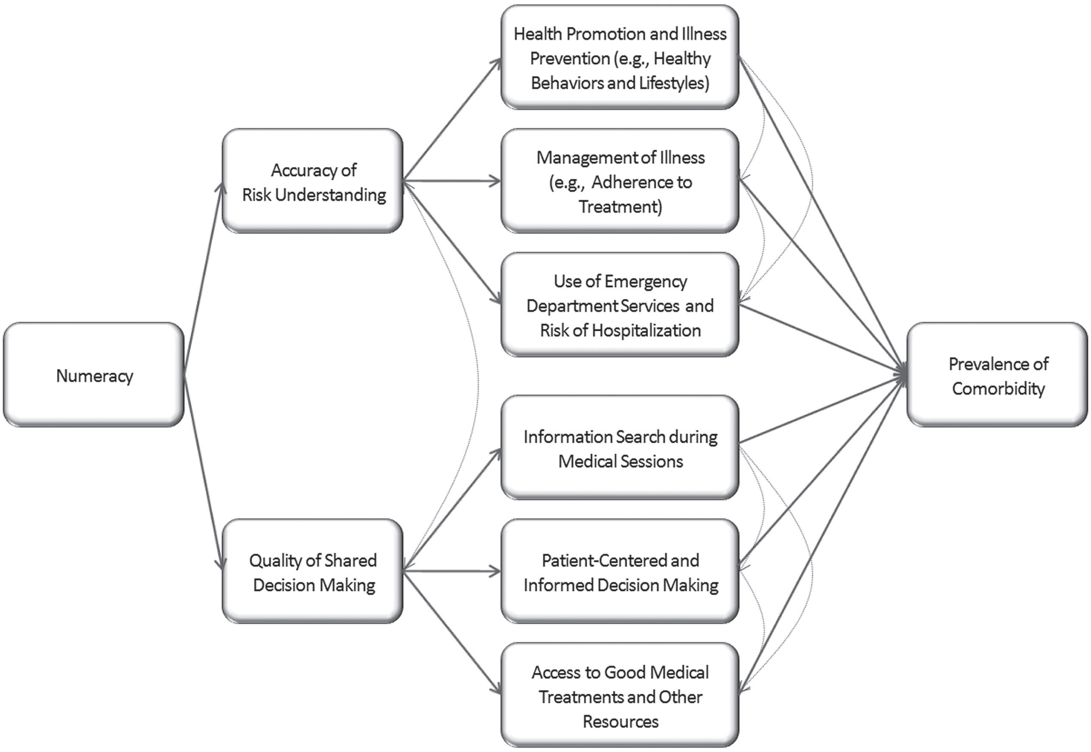

Figure 1. Conceptual framework of the effect of numeracy on prevalence of comorbidity and the mediating effect of accuracy of risk understanding and quality of shared decision making.

McAuliffe, & Nattinger, 2004; Gurmankin, Baron, & Armstrong, 2004), and they overestimate the efficacy of uncertain treatments (Weinfurt et al., 2003). Patients with low numeracy also have more difficulty interpreting the risks of side effects (Gardner, McMillan, Raynor, Woolf, & Knapp, 2011), they make less accurate diagnostic inferences based on numerical information about screening (Garcia-Retamero & Hoffrage, 2013; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2007; Petrova, Garcia-Retamero, Catena, & van der Pligt, 2016), and they are less able to use this information to adjust their risk estimates (Schwartz et al., 1997). Less numerate patients are also more easily biased by the way numerical information is framed (Garcia-Retamero & Galesic, 2010a, 2011; Peters & Levin, 2008; Peters et al., 2006). In part, these effects may emerge because less numerate patients are more influenced by non-numerical information, like affective mood states (Peters, Dieckmann, Dixon, Hibbard, & Mertz, 2007; Petrova, van der Pligt, & Garcia-Retamero, 2014).

The accuracy of perceptions of health-related benefits and risks may cascade, thereby affecting the prevalence of comorbidities by influencing patients' efforts to promote health and prevent disease, and efforts to comply with diagnosis and treatment directives (e.g., management of illness and hospitalization; see Figure 1; see also Paasche-Orlow & Wolf, 2007; von Wagner, Steptoe, Wolf, & Wardle, 2009). Disease prevention often depends on taking actions now to reduce risks and prevent uncertain consequences later. Still, patients with low numeracy act as if they value immediate rewards much more than temporally distant rewards (Nelson, Reyna, Fagerlin, Lipkus, & Peters, 2008)—presumably because they do not carefully evaluate or accurately interpret long-term probabilistic information and implications (Reyna et al., 2009). Patients with low numeracy also have difficulty weighing the risks and benefits of medical treatments and healthy behaviors (Fagerlin, Ubel, et al., 2007), and so they may be more resistant to generally effective interventions (e.g., education; Cavanaugh et al., 2008; Garcia-Retamero & Cokely, 2014a; Martin et al., 2012; Nelson et al., 2008).

Patients' ability to effectively manage their illness and mitigate health risks also often depends on the accuracy of risk understanding (Brewer et al., 2007). Patients with low numeracy can have difficulties understanding the probabilistic link between adherence and treatment effectiveness, which is consistent with evidence indicating that less numerate patients show lower medication compliance as compared with patients with more adequate numerical skills (Lindau, Basu, & Leitsch, 2006; Waldrop-Valverde et al., 2009, 2010). Inaccuracy of risk perceptions (e.g., about the likelihood of suffering a disease or the benefits of uncertain treatments) may also partially explain why patients with low numeracy use emergency department services more often and are at higher risk for hospitalization (Ginde et al., 2008). In short, these results suggest that numeracy may have indirect effects on prevalence of comorbidity through its effect on accuracy of risk understanding, which affects health promotion, health risk mitigation, and health-relevant planning.

Obstacles to shared decision making may also shape the effect of numeracy on prevalence of comorbidity (see Figure 1). Compared with patients with adequate levels of numeracy, less numerate patients often avoid asking doctors questions about their symptoms and medical treatments (Paasche-Orlow & Wolf, 2007), they spend less time gathering information about their disease during medical sessions (Portnoy, Roter, & Erby, 2010), and they use personal health records on the Internet less often (Sharit et al., 2014).

Less numerate patients also tend to favor a paternalistic model of medical decision making, in which doctors are dominant and autonomous and make decisions on their patients' behalf (Garcia-Retamero et al., 2014), whereas patients prefer not to participate and instead delegate decision making (Galesic & Garcia-Retamero, 2011b). This finding is troubling given that the paternalistic model of medical decision making has well-documented ethical and practical limitations (Kaplan & Frosch, 2005). Accordingly, patients with low numeracy are more likely to report negative interactions with their doctors (Manganello & Clayman, 2011; Roter, 2005), potentially influencing their subsequent information search and their ability to engage in decision making (McCaffery, Smith, & Wolf, 2010). These negative interactions also limit patients' access to good medical treatments and other health resources (e.g., regular medical check-ups, screenings, and immunization; Abdel-Kader et al., 2010; Hibbard, Peters, Dixon, & Tusler, 2007). Such obstacles to access of high-quality health care can ultimately translate into increased comorbidity risk.

In summary, there is a large and consistent body of research showing that many people have problems understanding simple, health-relevant numerical expressions of probability about health. This research indicates that many people are functionally innumerate, making them vulnerable regardless of other protective factors. In contrast, people who have well-developed numerical skills tend to have more accurate perceptions of health-related benefits and risks, and often make better decisions and plans that translate into important health benefits (e.g., lower prevalence of comorbidity). Educational efforts designed to improve risk understanding and decision-making skills are crucial elements of potential long-term solutions (Bruine de Bruin, Parker, & Fischhoff, 2007; Peters et al., 2014). Another essential element is the use of powerful, simple interventions that can have substantial benefits at minimal costs, particularly when designed to serve vulnerable populations with limited numeracy.

IMPROVING RISK UNDERSTANDING AND DECISION MAKING WITH VISUAL AIDS

Visual aids are simple graphical representations of numerical expressions of probability and include icon arrays, bar and line charts, and others (Ancker, Senathirajah, Kukafka, & Starren, 2006; Hildon, Allwood, & Black, 2012; Spiegelhalter, Pearson, & Short, 2011). Visual aids have long been known to confer benefits when communicating risk informa-

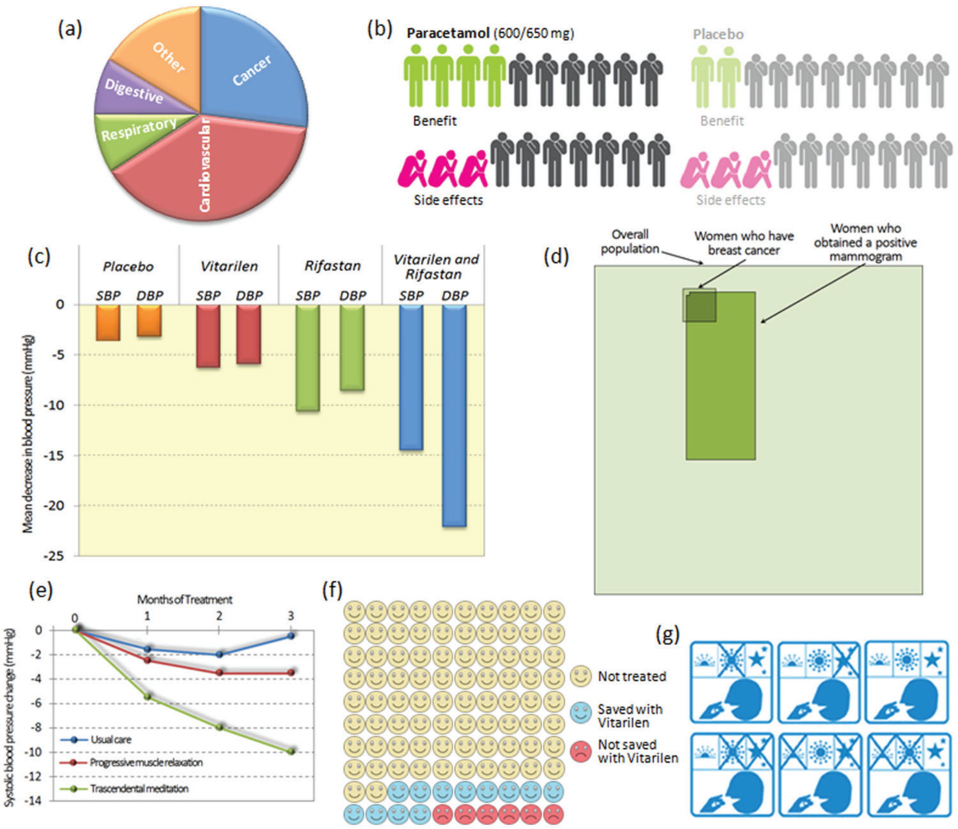

Figure 2. Examples of transparent visual aids. (a) A pie chart reporting the proportion of deaths by cause of death. (b) Icon arrays representing benefits and side effects of a medical treatment and a placebo. (c) A bar chart comparing the efficacy of two medical treatments (DBP = diastolic blood pressure; SBP = systolic blood pressure). (d) A visual grid to help infer the predictive value of mammography screening. (e) A line plot comparing the efficacy of several therapies. (f) Icon arrays to communicate treatment–risk reduction. (g) Pictograms reporting dosage, timing, and action information about prescribed medications. Adapted from "Communicating Health Risks With Visual Aids," by R. Garcia-Retamero and E. T. Cokely, 2013, *Current Directions in Psychological Science*, *22*, pp. 392–399. Copyright 2013 by The Authors. Adapted with permission.

tion about health (Edwards, Elwyn, & Mulley, 2002; Lipkus, 2007; Paling, 2003). However, not all visual aids are equally effective. Visual aids tend to provide an efficient means of risk communication when they are *transparent* (Garcia-Retamero & Cokely, 2013)—that is, when they promote representative (or unbiased) risk understanding and evaluation. Generally, this transparency means that the elements of the visual aid are well defined and they accurately and clearly represent the essential risk information by making part-to-whole relationships in the data visually available and comparable (see Figure 2 for some examples of transparent visual aids).

Previous research indicates that transparent visual aids can confer major decision-making benefits for various screenings, medical treatments, and lifestyle choices. Transparent visual aids typically improve the acquisition of both gist (general impression) and verbatim (precise) risk understanding (Feldman-Stewart, Brundage, & Zotov, 2007; Hawley et al., 2008; Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin, 2010a) by supporting evaluation of complex concepts, like incremental risks (Zikmund-Fisher et al., 2008); promoting consideration of beneficial treatments despite side effects (Feldman-Stewart, Kocovski, McConnell, Brundage, & Mackillop, 2000; Waters, Weinstein, Colditz, & Emmons, 2007); and informing patients' decisions about effective medical interventions and their influence on quality of life (Brundage et al., 2005).

Transparent visual aids also reduce errors caused by anecdotal narratives (e.g., treatment outcomes in unrepresentative patients; Fagerlin, Wang, & Ubel, 2005), they increase probability of health-promoting behaviors (e.g., vaccination, stopping smoking; Cox, Cox, Sturm, & Zimet, 2010; Garcia-Retamero & Cokely, 2011; Vogt & Marteau, 2012), they increase adherence to medical treatments in patients at risk (Dowse & Ehlers, 2005; Machtinger et al., 2007), and they improve recall of health-relevant information (Gaissmaier et al., 2012; Kakkilaya et al., 2011). Finally, transparent visual aids can increase risk avoidance (Schirillo & Stone, 2005; Stone, Gabard, Groves, & Lipkus, 2015; Stone, Yates, & Parker, 1997), they can be more beneficial than other transparent information formats (e.g., fact boxes or tables; Hawley et al., 2008; Petrova, Garcia-Retamero, & Cokely, 2015), and they are often perceived as easier to understand than numbers (Goodyear-Smith et al., 2008). Unfortunately, the benefits of transparent visual aids can be different for different people; one size does not necessarily fit all.

THE INFLUENCE OF INDIVIDUAL DIFFERENCES ON THE EFFICACY OF VISUAL AIDS: A SYSTEMATIC REVIEW

Following the recommendations of the Cochrane handbook for systematic reviews (Higgins & Green, 2008), we conducted a systematic review of research investigating the influence of skills on the efficacy of visual aids. We focused on individual differences in two relevant skills that influence the efficacy of visual aids (Garcia-Retamero & Cokely, 2013, 2014b)—that is, numeracy and graph literacy. As noted earlier, numeracy is broadly defined as one's practical ability to use mathematical skills to solve everyday problems, and statistical numeracy in particular has been found to be one of the strongest single predictors of general decision-making skill and risk literacy (Cokely et al., 2012; see also Ancker & Kaufman, 2007; Anderson & Schulkin, 2014a; Peters, 2012; Peters & Bjalkebring, 2015; Reyna et al., 2009). Broadly, graph literacy is the ability to evaluate and extract data and meaning from graphical representations of numerical information, which is another essential component of risk literacy (Garcia-Retamero, Petrova, Feltz, & Cokely, in press).

Our search first targeted articles written in English and published in scientific journals between January 1995 and April 2016. We searched the databases Web of Science, PubMed, PsycINFO, ERIC, Medline, and Google Scholar, using combinations of the following keywords: *numeracy*, *risk literacy*, or *health literacy* and *visual aid* or *graph*. We identified additional studies from the reference list of the articles selected and via contacts with allied professionals and experts. We also searched the gray literature (e.g., unpublished studies and congress abstracts). The initial search returned 440 publications.

We then selected articles based on the following criteria: (a) They included empirical studies (i.e., we excluded review papers), (b) they systematically investigated the effect of skills (i.e., we excluded articles that only controlled for their effect or do not report results about the effect of skills), and (c) they investigated the effect of visual aids on objective measures (e.g., risk understanding, decision making, and healthy behavior) or subjective measures (e.g., preferences and opinions), or they compared their effect on these measures with that of numerical information reported in text (i.e., we did not review research comparing the efficacy of visual aids with that of other formats, such as fact boxes or tables). A total of 36 publications were identified for inclusion.

For each paper, we recorded the following information: (a) demographics (i.e., sample size, type of participants, gender, average age, education achievement, and nationality), (b) measure of skills (i.e., whether the authors investigated the effect of numeracy and the instrument[s] that they used to measure this skill, whether the authors investigated the effect of graph literacy and the instrument[s] that they used to measure this skill, and whether the authors investigated the interaction between numeracy and graph lit-

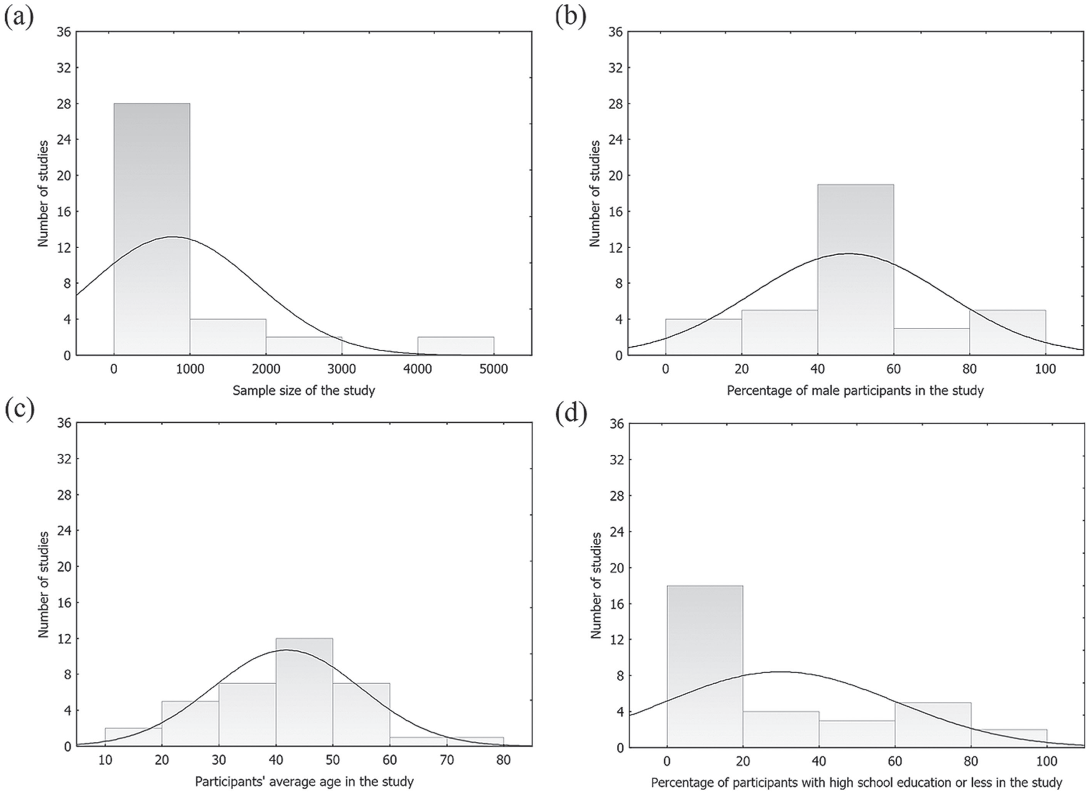

Figure 3. Demographic characteristics of participants in the studies included in the systematic review. (a) Distribution of the sample size (*n* = 36). (b) Distribution of the percentage of male participants (*n* = 36). (c) Distribution of participants' average age (*n* = 35). (d) Distribution of the percentage of participants with high school education or less (*n* = 33).

eracy), (c) manipulation of information format (i.e., type of visual aids used and whether the authors included a control condition reporting risks as numbers in written text), (d) dependent variable(s) (i.e., type of measures included in the study), (e) whether the authors investigated psychological mechanism(s) explaining the effect of visual aids and type of mechanism(s) investigated, (f) results in objective measures, (g) results in subjective measures, and (h) whether there is a match between results in objective and subjective measures. Tables A1 to A3 in the appendix report a summary of the basic characteristics of the studies and the main results.

Description of the Studies

The reviewed studies sampled a total of 27,885 participants. The average sample size across studies is 775 (range = 47–4,685), and the distribution of the number of participants in the studies is positively skewed (see Figure 3). Gender (55.8% male) and age (average = 45.02, range = 17–94) were roughly balanced. We observed a wide range of educational achievement, with 37% of participants having a bachelor's or higher college degree, 34% of participants having some college education, and 29% having completed high school or less education.

The studies involved demographically diverse samples of participants, including probabilistic national (i.e., representative) samples (8%), general community samples (11%), online panels (31%), patients and high-risk individuals (22%), practicing physicians (6%), highly educated young adults (25%), and older adults with limited skills (3%). Overall, the studies involved participants from 60 different countries from all over the world; 14% of the studies were cross-cultural (i.e., they were conducted in more than one country), and the rest were conducted in only one country, including the United States (50%), Spain (11%), Germany (8%), and Switzerland (8%), among others.

Ninety-seven percent of the reviewed studies measured numeracy, 33% measured graph literacy, and 31% measured both. Sixty-one percent of the studies that measured numeracy used objective tests: 86% of these studies used the test developed by Lipkus et al. (2001) and/or Schwartz et al. (1997) described earlier, and 23% used the Berlin Numeracy Test (Cokely et al., 2012)—an efficient test designed to measure numeracy and predict risk literacy in broad and highly educated samples. The instrument is internally consistent, predicting answers to ecologically valid probabilistic medical decisions better than a wide range of other ability and numeracy tests (Cokely et al., in press; Cokely, Ghazal, & Garcia-Retamero, 2014). In addition, 39% of the studies that measured numeracy used the subjective numeracy scale developed by Fagerlin, Zikmund-Fisher, et al. (2007). This scale measures perceptions of numerical ability and preferences for numeric information presentation. This scale is significantly correlated with scales measuring objective numeracy (Galesic & Garcia-Retamero, 2010), and it predicts objective performance in problems involving numerical risks (Zikmund-Fisher, Smith, Ubel, & Fagerlin, 2007).

All (100%) of the studies that measured graph literacy used the objective graph literacy scale developed by Galesic and Garcia-Retamero (2011c). The scale measures three abilities of graph comprehension: (a) the ability to read the data, that is, to find specific information in the graph (e.g., the ability to read off the height of a particular bar within a bar chart); (b) the ability to read between the data, that is, to find relationships in the data as shown on the graph (e.g., the ability to read off the difference between two bars); and (c) the ability to read beyond the data, or make inferences and predictions from the data (e.g., the ability to project a future trend from a line chart). The scale is generally a robust predictor with good psychometric properties, and it was validated in several samples from different cultures (Garcia-Retamero & Muñoz, 2013; Sharit et al., 2014; Woller-Carter, Okan, Cokely, & Garcia-Retamero, 2012; Zikmund-Fisher, Witteman, et al., 2014; see Garcia-Retamero et al., in press, for a review). In addition, one study also used the subjective graph literacy scale developed by Garcia-Retamero, Cokely, Ghazal, and Joeris (2016)—a brief and reliable instrument that predicts graph comprehension by measuring people's judgment about their own ability to interpret graphs. The subjective graph literacy scale achieves a high level of psychometric performance largely comparable to that of the objective graph literacy scale while limiting test anxiety. The subjective graph literacy scale also takes less than 10% of the time of the objective test (i.e., less than 1 min on average), which makes it ideal for clinical research and practice.

Finally, 97% of the studies used objective measures of performance as dependent variables, 61% used subjective measures, and 58% used both. The studies on objective performance measured accuracy of risk understanding and/or knowledge about risk (e.g., verbatim or gist; 72%), accuracy of decision making (19%), accuracy of recall (14%), judgment calibration (i.e., accuracy of self-assessments of accuracy of risk understanding; 6%), reading and estimate latency (8%), and eye movements when viewing graphs (e.g., amount of time spent when watching areas of interest in graphs; 8%). In the studies that used subjective measures, authors examined opinions about accessibility of the information (e.g., comprehensibility, usefulness, and difficulty of related questions; 31%), opinions about attractiveness of the information (e.g., overall impression, attractiveness of colors, imagery, technical implementation; 8%), preferences of information format (e.g., graph preferences; 22%), familiarity with information format (3%), ease of imagining risk (3%), confidence (11%), self-assessments of accuracy of risk understanding (3%), information trust (11%), attitudes and affective reactions toward the information provided (11%), and behavioral intentions toward a behavior (14%). Perceptions of risks were measured in 25% of the studies. These perceptions were considered as objective measures when authors (a) computed correlations with correct values and/or (b) investigated whether participants differentiate between different probability levels (22% of the studies).

IMPROVING RISK LITERACY WITH VISUAL AIDS

In most (83%) of the reviewed studies, the authors investigated whether static (i.e., nonanimated) visual aids improve risk literacy and/or promote healthy behavior; 87% of these studies involved the effect of static visual aids in people with different levels of numeracy; 65% of these studies included a control condition (i.e., they investigated the efficacy of static visual aids as compared with presenting numerical information in written text), and the rest of the studies did not include a control condition.

Eighty-eight percent of the studies showed that static visual aids tend to be beneficial, and only 12% failed to detect notable benefits of visual aids in the evaluated tasks. These percentages excluded (a) studies including distorted graphs or graphs depicting only number of people at risk as they were not designed to improve risk understanding, and (b) studies that did not report results in objective measures (e.g., they reported only preferences and/or opinions). Most (80%) of the studies showed that static visual aids tend to have some beneficial effect for people with low numeracy (vs. 73% for those with high numeracy) when comparing the effect of visual aids with that of numbers in text, and two thirds (67%) of the studies showed larger improvements in risk literacy among people with low numeracy (vs. 33% showing smaller or no improvement). To the extent that these studies are ecologically valid and representative, the current results indicate that static visual aids tend to cause larger improvements in risk literacy and decision making among less numerate individuals than among more numerate participants. Nevertheless, 83% of the reviewed studies indicated that less numerate individuals do not generally achieve as high a level of performance as those with higher numeracy. People with high numeracy tend to be more accurate overall. All in all, static visual aids are helpful, but many people with low numeracy might still have some difficulties interpreting and using numerical concepts when they are visually represented.

Regardless of their numerical skills, people also differ in their ability to understand graphically presented quantitative information (i.e., graph literacy). In 37% of the reviewed studies, the authors investigated the benefits of static visual aids in people with different levels of graph literacy, 45% of which also included a control condition. All (i.e., 100%) of these studies showed that static visual aids were especially helpful for people with at least a minimum level of graph literacy. In addition, authors of three studies (10%) investigated whether people with different levels of numeracy *and* graph literacy differed in the extent to which they profited from static visual aids. These authors investigated the interactive effect of the two factors, revealing the essential role of graph literacy in visual aid comprehension: Static visual aids were particularly helpful for people with low numeracy so long as they also had moderate-to-high graph literacy. Not surprisingly, however, people who could not accurately interpret or evaluate graphs typically did not benefit from visual aids regardless of their overall level of numeracy.

To further illustrate the essential performance cut-off that results when people fall below a minimum level of graph literacy, consider a study conducted by Garcia-Retamero and Galesic (2010b). The authors examined the accuracy of perceptions of the effectiveness of a medical treatment in probabilistic national samples in the United States and Germany. They compared the efficacy of different types of static visual aids (i.e., icon arrays and bar graphs), representing either the entire population at risk or affected individuals only. In addition, the authors tested the efficacy of these visual aids when the numerical information added to the visual aids was presented as either absolute or relative risk reduction.

Garcia-Retamero and Galesic (2010b) observed similar increases in judgment and decision accuracy when using icon arrays and bar graphs. Static visual aids were found to be useful additions when the numerical information was presented either in terms of relative or absolute risk

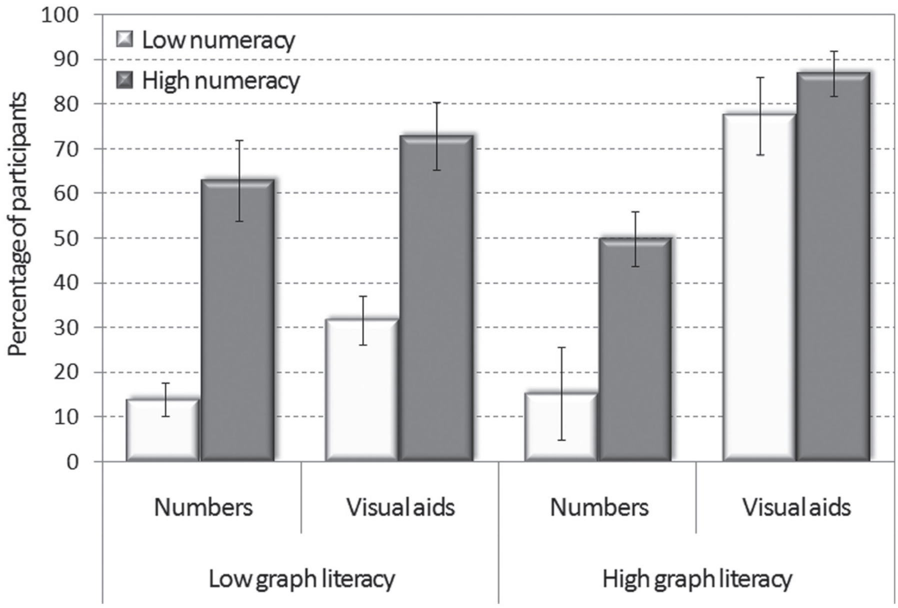

Figure 4. Percentage of participants with low and high numeracy and low and high graph literacy who correctly inferred treatment risk reduction, as a function of condition (numbers vs. visual aids). In the visual-aids condition, icon arrays and bars reported the entire population at risk. Error bars represent one standard error. Adapted from "Who Profits From Visual Aids: Overcoming Challenges in People's Understanding of Risks," by R. Garcia-Retamero and M. Galesic, 2010, *Social Science & Medicine*, *70*, 1019–1025. Copyright 2010 by Elsevier Ltd. Adapted with permission.

reductions. Importantly, these visual aids were most beneficial for individuals who had low numeracy but at least a modest level of graph literacy—especially when the visual aids depicted the entire population at risk (see Figure 4). Among this group of innumerate yet moderately graph-literate people, accuracy increased from less than 20% to nearly 80% when visual aids were used. The benefits of transparency were so large and robust that providing visual aids fully eliminated differences in accuracy between highly numerate and functionally innumerate people, so long as the less numerate individuals were also graph literate.

As expected, Garcia-Retamero and Galesic (2010b) showed that static visual aids were not beneficial for people who had both low numeracy and low graph literacy. These results are consistent with those in a study conducted by Ruiz et al. (2013) in patients with low numeracy and graph literacy in the United States, showing that visual aids tend to confuse rather than help these patients (see also Garcia-Retamero & Muñoz, 2013, and Nayak et al., 2016). Authors of other studies investigated whether the benefits of visual aids extend beyond risk understanding and risk estimate accuracy.

Does Accuracy of Risk Understanding Translate Into Better Decision Making?

Visual aids that promote risk understanding also tend to improve decision making, promoting healthy behavior, especially in people with low numeracy and moderate-to-high graph literacy. In a series of longitudinal studies, Garcia-Retamero and Cokely (2011, 2014a, 2015b) examined the effects of a risk awareness intervention in a large sample of high-risk individuals in Spain (i.e., sexually active young adults with limited numerical skills). The authors manipulated the formats used to design health messages. In particular, half the participants received a message focused on the benefits afforded by adopting a behavior (a gain-framed appeal); the rest of the participants received a message focused on the costs associated with failing to adopt the behavior (a loss-framed appeal). In accord with the leading theory (Rothman, Martino, Bedell, Detweiler, & Salovey, 1999), the authors showed that gain-framed messages induced greater adherence for a prevention behavior (condom use). In contrast, loss-framed messages were more effective for promoting an illness-detecting behavior (screening for sexually transmitted diseases). Importantly, these results held controlling for message length and general content. Results also held when accurate numerical information about sexually transmitted diseases was included in the messages.

Nevertheless, when static icon arrays were added to the health information, the framing bias was eliminated. Both the gain- and loss-framed messages became equally and highly effective for promoting condom use and screening, increasing healthy behavior. In addition, young adults with low numeracy benefited more from the use of static icon arrays than those with higher numeracy as long as they were moderately-to-highly graph literate—a result that is consistent with those in the studies reported earlier (Garcia-Retamero & Galesic, 2010; Garcia-Retamero & Muñoz, 2013; Nayak et al., 2016). Finally, the authors showed that static icon arrays were as effective for promoting prevention of sexually transmitted diseases (i.e., condom use) as an extensive 8- to 10-hr evidence-based educational risk awareness program (see Garcia-Retamero & Cokely, 2015a, for a review).

These results accord with others showing that static icon arrays and bar graphs can improve the accuracy of decisions about medical treatments and risky prospect evaluations in people with low and high numeracy by directly improving risk understanding (Hawley et al., 2008; Oudhoff & Timmermans, 2015; see also Garcia-Retamero & Dhami, 2013). Unfortunately, Ruiz et al. (2013) again showed that static icon arrays fail to provide benefits when people do not have at least a minimum level of graph literacy.

Does the Format Matter? On the Effect of the Type of Visual Aids

In almost half (i.e., 46%) of the studies included in our systematic review, authors investigated the benefits of only one type of visual aids; 75% of these studies focused on static icon arrays, and the rest involved other types of visual aids (e.g., static bar charts and grids). In contrast, authors of 37% of the studies compared the efficacy of different types of static visual aids (e.g., icon arrays, bar and line charts, grids, and pies), or they compared different types of static icon arrays. Finally, in 17% of the studies, the authors investigated the effect of dynamic visual aids (i.e., animated graphs with dynamic design features that unfold over time) or interactive visual aids (e.g., grids of squares that reveal stick figures when clicked). These authors examined whether the optimal type of visual aids depends on the level of numeracy and/or graph literacy, revealing mixed results.

In 92% of the studies on the effect of static icon arrays that did not compare the effect of these visual aids with that of others, researchers concluded that static icon arrays could improve accuracy of risk understanding and recall. In addition, 75% of the studies comparing the effect of this type of visual aids with that of numbers reported in text showed that static icon arrays tended to be particularly helpful for people with low numeracy (i.e., these people show larger improvements in accuracy than those with high numeracy). The benefits of icon arrays partially stemmed from the ability to leverage basic perceptual features of these displays, which are among the easiest features to process (see also Price, Cameron, & Butow, 2007).

In contrast, only 14% of the studies that compared different types of static visual aids suggested that static icon arrays were more effective than other types of visual aids; 29% of these studies suggested that static icon arrays tend to be less helpful, yet 57% suggested that these icon arrays are as effective as other types of visual aids (e.g., bar charts). For instance, Hamstra et al. (2015) examined accuracy of understanding of longitudinal risks (i.e., at multiple time points) in patients in the United States. The authors depicted risks using static icon arrays, pies, and line and bar graphs. Results showed that static icon arrays and bar graphs were the most effective means for improving understanding of longitudinal risks in patients with either low or high numeracy, eliminating differences in accuracy between the two groups—a result in

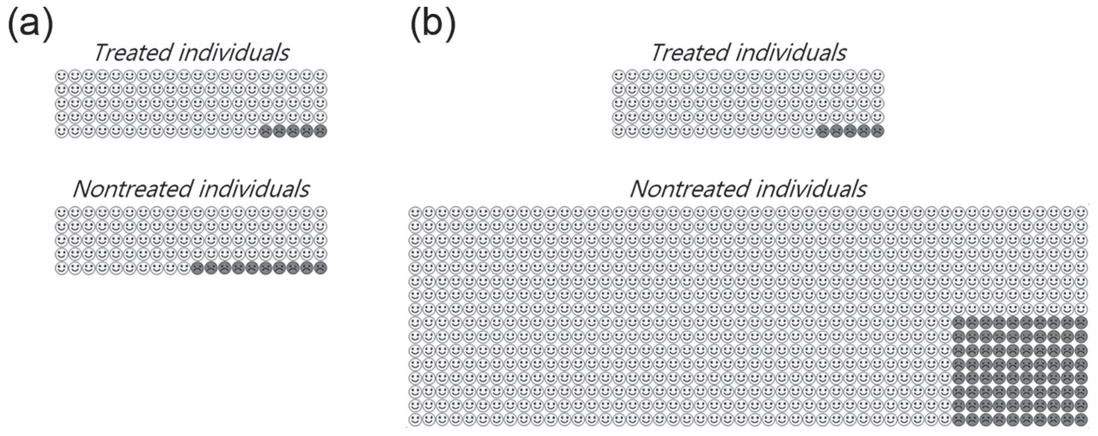

Figure 5. Icon arrays representing a treatment risk reduction of 50% with (a) equal and (b) different samples of treated and nontreated individuals. Affected individuals are represented in dark gray. Healthy individuals are represented in light gray.

line with those in the study by Garcia-Retamero and Galesic (2010b) described earlier.

Based on these and other results, our analysis of the available studies suggests that not all types of static icon arrays are equally helpful for people with different levels of numeracy and graph literacy. We speculate that these differences reflect structural differences in the tasks and the quality of the design of the visual aid. For example, icons can be randomly scattered throughout the icon array or can be arranged systematically in a block. Research shows that static icon arrays with scattered icons perform poorly compared with static icon arrays in which icons are systematically arranged (Feldman-Stewart et al., 2007). Scattered icon arrays are also particularly difficult for people with low numeracy, who rated them as the least preferred information format (Ancker, Weber, & Kukafka, 2011a; Zikmund-Fisher et al., 2012). Accordingly, scattered icon arrays are generally not recommended to convey quantitative information. However, the unique potential benefit of these arrays is that they can convey the sense of randomness at the same time as conveying the quantity, suggesting they may be useful for some very specialized types of risk communications (Feldman-Stewart et al., 2007).

In addition, research shows that visually depicting both the number of individuals affected by a risk (i.e., the number of times a target event happens, or the numerator) and the overall number of individuals at risk (i.e., the overall number of opportunities for the event to happen, or the denominator) improves accuracy of risk understanding (Bruine de Bruin, Stone, Gibson, Fischbeck, & Shoraka, 2013). This is usually the case for people with low numeracy and/or high graph literacy (Garcia-Retamero & Galesic, 2010b). In contrast, depicting only the numerator increases perceptions of risk and risk aversion, and decreases risk understanding in people with low and high numeracy alike (Stone et al., 2015). The benefits of representing both numerators and denominators in icon arrays stem from visually depicting the part-to-whole ratio of the risk magnitude (Ancker et al., 2006), thereby making the part-to-whole relationship "pop out" (Reyna, 2008; Reyna & Brainerd, 2008). Adding numerical and textual information describing the information depicted in the static icon arrays (i.e., numerators and denominators) also improves performance and reduces differences in accuracy between people with low and high numeracy (Hamstra et al., 2015; see also Feldman-Stewart et al., 2007).

Even if depicting both numerators and denominators is broadly effective, recent research suggests that many people show denominator neglect. That is, they tend to focus more on numerators and pay less attention to denominators (Denes-Raj, Epstein, & Cole, 1995; Reyna, 2004; Stone et al., 2003). This bias is particularly problematic for people with low numeracy when comparing risks with different sample sizes. To illustrate, when estimating the effectiveness of medical treatments, people often receive information about the number of affected individuals in the treated and nontreated groups. This information can be reported using samples of treated and nontreated individuals of the same size (same denominators; e.g., a group of 100 individuals who receive a treatment and a group of 100 individuals who do not receive the treatment or receive a placebo; see Figure 5). In such a case, people can infer treatment risk reduction by focusing on absolute numbers of affected individuals (numerators; e.g., number of sick individuals in each group, or 5 vs. 10 in Figure 5). In contrast, the information can also be reported using samples of treated and nontreated individuals of different size (different denominators; e.g., a group of 100 individuals who receive a treatment and a group of 800 individuals who do not receive the treatment or receive a placebo). In such a case, people would need to compute or otherwise represent proportions of individuals in the groups to accurately infer treatment risk reduction (5 in 100 vs. 80 in 800 in Figure 5)—which is something less numerate people generally fail to do (Garcia-Retamero & Galesic, 2009; Garcia-Retamero, Galesic, & Gigerenzer, 2010).

Depicting the risks using icon arrays can effectively reduce or eliminate denominator neglect, especially in people with low numeracy and/or high graph literacy (Garcia-Retamero & Galesic, 2009; Garcia-Retamero et al., 2010). Nevertheless when comparing different risks, it is recommended that risks be depicted using icon arrays with equal denominators, as icon arrays with different denominators reduce but do not completely eliminate the bias in people with low graph literacy (Garcia-Retamero, Cokely, Ghazal, et al., 2016; Okan, Garcia-Retamero, Cokely, & Maldonado, 2012, 2015).

Icon arrays also differ in iconicity, or the extent to which a representation resembles what it is supposed to represent (Moles, 1966). Icon arrays can include concrete, anthropomorphic icons (e.g., gender-matched restroom person icons, actual head-and-shoulder photographs of people of varied races, smile/frown faces), or they can be more abstract (e.g., they can include filled circles). In two studies, authors investigated whether the level of iconicity of static icon arrays affects comprehension and recall in people with different levels of numeracy and/or graph literacy. In one of the studies, Gaissmaier et al. (2012) investigated accuracy of understanding of benefits and side effects of medical interventions in young adults in Germany. The authors designed different representations of health-related statistical information ranging from very low iconicity (black-and-white icon arrays including rectangular blocks) to very high iconicity (icon arrays including color photographs). The authors compared the efficacy of these visual aids with that of numbers in text. The authors concluded that the actual level of iconicity of the static icon arrays did not matter. Instead, the only difference between representations that affected comprehension and recall was the difference between numbers and icon arrays. Individuals with high graph literacy had better comprehension and recall when presented with icon arrays instead of numbers, whereas the reverse was true for individuals with low graph literacy.

These findings contrast with those in a study conducted by Zikmund-Fisher, Witteman, et al. (2014) in a diverse sample of people from the general population in the United States. The authors investigated perceptions of risk, recall, and opinions about static icon arrays that also differed in iconicity. The authors showed that the type of icon array influenced risk perceptions and recall, and the optimal type of icon array depended on people's level of numeracy and graph literacy. In particular, risk perceptions were more accurate (i.e., they showed larger correlations with actual values) among people with high numeracy or high graph literacy when the risks were depicted using anthropomorphic icons (e.g., static gender-tailored restroom icons), resulting in better outcomes. In contrast, these icon arrays did not perform better than other types of icon arrays among people with low numeracy or low graph literacy. Risks depicted using restroom icons were also easier to recall, especially for people with high numeracy or high graph literacy, although they also tended to improve recall in people with low numeracy or low graph literacy. In addition, participants with low and high numeracy and/or graph literacy rated the personlike icon arrays as their most preferred information format.

There are several plausible hypotheses for differences between the two studies on the influence of iconicity of icon arrays. First, only the study by Gaissmaier et al. (2012) investigated the efficacy of icon arrays as compared with presenting numerical information in written text (i.e., they included a control condition), which may have attenuated differences in performance between the conditions including visual aids (vs. control). Second, in the study by Zikmund-Fisher, Witteman, et al. (2014) but not in the Gaissmaier et al. study, the anthropomorphic icon arrays were gender tailored to the participant and reported actual personal risk, which may have increased attention or perception of the information as personally relevant. Finally, and perhaps most important, Gaissmaier et al. assessed accuracy of risk understanding (i.e., knowledge) as a primary outcome, whereas the primary outcome in the study by Zikmund-Fisher, Witteman, et al. was perceptions of risk. Although these authors reported correlations between these perceptions and actual values, it may be that the type of icon array makes less of a difference in comprehension than in more subjective responses to the risk. Taken together, it seems that iconicity might not influence objective risk understanding (Gaissmaier et al., 2012), although it may affect more subjective experiences and responses to risk (Zikmund-Fisher, Witteman, et al., 2014). Authors of future research should investigate this hypothesis in greater detail.

Finally, several studies compared the efficacy of interactive and dynamic icon arrays with that of static icon arrays in people with different levels of numeracy and/or graph literacy. Most (83%) of these studies converge to show that interactive and dynamic design features do not necessarily contribute to improve risk comprehension and decision making in these people. For instance, Zikmund-Fisher, Dickson, and Witteman (2011) showed that providing people with numerical information and requiring them to depict this information interactively in an icon array did not have a beneficial effect. In their study, icon arrays were initially blank, and participants had to use their mouse to click and drag in the icon arrays to set them to the appropriate risk level. The icon arrays continually adjusted to provide feedback. Participants using interactive icon arrays made less accurate inferences and poorer decisions than those who viewed the information displayed in equivalent static icon arrays. Other authors showed that even if interactive icon arrays do not improve risk perceptions or recall, they can reduce disparities in emotional reactions toward risk between people with low and high numeracy (Ancker, Weber, & Kukafka, 2011b; Mason et al., 2014). Interactive icon arrays are also perceived as visually appealing (Ancker et al., 2011a).

Studies on the effect of dynamic icon arrays showed both a decline in performance and poorer evaluations with the addition of the animation to static icon arrays. To illustrate, Zikmund-Fisher et al. (2012) presented participants flash-based animated icon arrays that (a) represented risk events one at a time, (b) had scattered icons settled into a group, or (c) had scattered icons shuffled themselves (either automatically or by participants' control). The authors concluded that dynamic icon arrays including these features did not improve risk understanding or recall, and some (e.g., dynamic displays with scattered icons) performed particularly poorly in people with high numeracy. Static arrays that grouped icons, however, performed well on measures of knowledge and decision accuracy (see Tait, Voepel-Lewis, Brennan-Martinez, McGonegal, & Levine, 2012, for similar results). In addition, people with low and high numeracy both reported the highest evaluation ratings when they received the static icon arrays. The novelty of some interactive and dynamic displays or the potential of some of their features to capture attention might have distracted participants, preventing encoding and processing of more essential information in the icon arrays (e.g., to represent part-to-whole relations in the data). In these ways, interactive and dynamic icon arrays can both add to people's burden and distract them from understanding relevant statistical information.

Results in the studies conducted by Zikmund-Fisher et al. (2011, 2012) conflict with those in a recent study conducted by Okan et al. (2015). The newer study found that simple dynamic design features can and do often enhance the efficacy and reach of static visual aids as long as they are designed to enhance (a) specific cognitive processes involved in graph comprehension (i.e., encoding and integration of information; Carpenter & Shah, 1998; Huestegge & Philipp, 2011) and (b) active elaborative processing and deliberation (Cokely, Kelley, & Gilchrist, 2006). Interestingly, dynamic icon arrays with these features do appear to improve performance in people with low graph literacy, even though these individuals typically experience only limited benefits from the use of static visual aids. Specifically, Okan et al. developed and tested three types of dynamic features: (a) the sequential presentation of the different elements of icon arrays, (b) the use of explanatory labels indicating what was depicted in the different regions of the arrays, and (c) the use of a reflective question followed by accuracy feedback. The first type of icon arrays (sequential icon arrays) was developed to encourage the allocation of attention and encoding of *all* regions of the displays. The second type of icon arrays (labeling icon arrays) was developed to encourage the allocation of attention to the elements that enable people to identify referents in the displays as well as to promote the integration of such information with that depicted in the displays. The third type of icon arrays (the transfer estimate icon arrays) was constructed to promote active, elaborative processing of the depicted information.

Okan et al. (2015) showed that explanatory labels improved risk understanding among less-graph-literate participants, whereas reflective questions resulted in large and robust performance benefits among participants with either low or high graph literacy. Participants who received the dynamic icon arrays with explanatory label or reflective questions made more accurate inferences than those who viewed the information displayed in equivalent static icon arrays or dynamic icon arrays that represent the information sequentially. Broadly, these results suggest that dynamic icon arrays designed to combine facilitation of the identification of referents (e.g., through the use of explanatory labels) with the promotion of active, elaborative processing (e.g., through the use of reflective questions) can enhance risk understanding and support informed decision making even among less-graph-literate individuals. In theory, this follows because these dynamic visual aids were also designed with a transparency standard in mind—that is, the design goal was to support the development of a more elaborate yet unbiased mental model of associated risks and trade-offs.

In summary, our systematic review of the research indicates that the type of visual aid used can affect accuracy of risk understanding in people with different levels of numeracy and graph literacy. Some of the studies that we reviewed in this section suggested that tailoring the type of visual aid to people's skills and the type of information needed generally increases accuracy of risk understanding and decision making (Hamstra et al., 2015; Hawley et al., 2008; Okan et al., 2015; Zikmund-Fisher, Witteman, et al., 2014). However, other studies showed that identifying the goal of the communication helps inform the selection of the best type of visual aid. For instance, consistent with long-standing human factors guidelines, several authors recommended using bar graphs to compare different data points, using line graphs to depict trends over time (Fischhoff, Brewer, & Downs, 2012; Lipkus, 2007; Lipkus & Hollands, 1999), using pie graphs to communicate information about proportions (Fischhoff et al., 2012), using grids to depict very large numbers (Garcia-Retamero, Cokely, & Hoffrage, 2015; Garcia-Retamero & Hoffrage, 2013), and using magnifier risk scales (including magnifying lenses) to depict very small numbers (Ancker et al., 2006). Unfortunately, tailoring the type of visual aid to people's skills and the type of task is not feasible in some contexts, particularly given that the incremental benefits in accuracy and decision making can sometimes be modest (Gaissmaier et al., 2012; Garcia-Retamero & Galesic, 2010b; Hamstra et al., 2015).

One general heuristic that appears broadly robust and has emerged across studies on features of effective visual aids is that, on average, "less is more" (Hess, Visschers, & Siegrist, 2011a; Hess, Visschers, Siegrist, & Keller, 2011; Zikmund-Fisher, Fagerlin, & Ubel, 2010; see also Peters et al., 2007; Shah & Hoeffner, 2002). For example, very simple icon arrays including clear explanations to convey the meaning of important information can improve risk under-

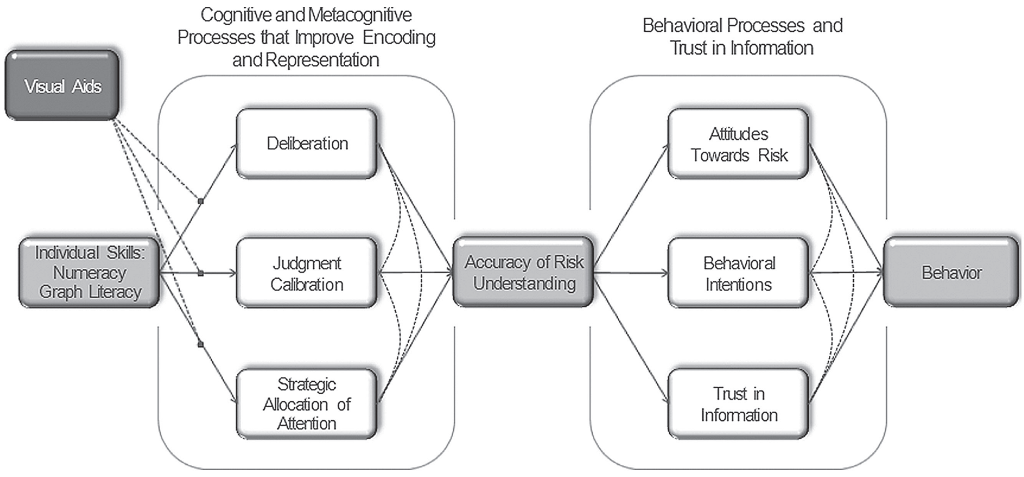

Figure 6. Conceptual framework of the psychological mechanisms explaining the effect of visual aids. The framework assumes that visual aids act as moderators of the effect of individual differences in numeracy and graph literacy on healthy behavior. This effect is mediated by (a) cognitive and metacognitive processes, (b) accuracy of risk understanding, and (c) behavioral processes and information trust.

standing (Garcia-Retamero & Cokely, 2014b; Garcia-Retamero & Galesic, 2010b; Okan et al., 2015; Zikmund-Fisher, Witteman, et al., 2014). These icon arrays offer a relatively efficient means of reaching individuals with different levels of numeracy and graph literacy. In addition, a simple training in the use of icon arrays (e.g., including reflective questions followed by accuracy feedback to encourage active, elaborative information processing) maximizes potential benefits and reaches vulnerable groups of people with limited graph literacy (Okan et al., 2015). Finally, our analysis also revealed that people with different levels of numeracy and graph literacy like, trust, and prefer simple icon arrays (see also Hawley et al., 2008; Zikmund-Fisher et al., 2010, 2012; Zikmund-Fisher, Witteman, et al. 2014). These icon arrays might then be the recommended information format if tailoring the type of visual aid to people's skills and the goal of the task is not feasible.

COGNITIVE AND BEHAVIORAL MECHANISMS EXPLAINING THE EFFICACY OF VISUAL AIDS

In 25% of the reviewed studies, authors investigated underlying mechanisms explaining the beneficial effects of visual aids on risk understanding and decision making for people with different levels of numeracy and/or graph literacy. Figure 6 depicts a conceptual framework based on our analysis, which can be broadly characterized by an indirect effect of numeracy and graph literacy on behavior through (a) cognitive and metacognitive processes (e.g., deliberation, judgment calibration, and strategy use), (b) risk understanding (e.g., an integrated situation model in long-term memory), and (c) behavioral processes (e.g., attitudes toward risk, behavioral intentions) and information trust. The framework assumes that visual aids act as moderators of the effect of individual differences in numeracy and graph literacy on accuracy of risk understanding and healthy behavior.

According to skilled decision theory (Cokely et al., in press), superior decision making under risk generally follows from heuristic deliberation (e.g., cognitive monitoring and control processes) deployed in the service of knowledge acquisition and evaluation (e.g., representative risk understanding). Consistent with this theory, research shows that visual aids improve accuracy of risk understanding in part because they increase the likelihood that people deliberate more elaboratively about the relevant risks and trade-offs. This idea was illustrated in a study conducted by Garcia-Retamero, Cokely, Wicki, and Joeris (2016) in a large sample of surgeons from 60 countries. Surgeons read a scenario describing the results of a randomized controlled trial examining the risk of postsurgical side effects in patients, and they provided an estimate of the risk. Half of the surgeons received the information in numbers; the other half received the information represented visually via a static icon array. Results showed that many surgeons were not numerate enough to correctly interpret the surgical risk without additional support. However, the intervention using visual aids enhanced risk comprehension, eliminating differences between surgeons with low and high numeracy.

Notably, the analyses conducted by Garcia-Retamero, Cokely, Wicki, et al. (2016) showed that visual aids improved risk comprehension among less numerate surgeons because visual aids increased the time that these surgeons spent deliberating about the risk-relevant information. That is, when the visual aids were used, more and less numerate surgeons tended to take about the same amount of time to make their risk estimates, reflecting more effective exploration, identification, and evaluation of essential aspects of the information about the risk. In contrast, when visual aids were not present, less numerate surgeons tended to ignore details of the information, perhaps because they could not interpret the information or because they erroneously assumed they understood the relevant details. Thus deliberation tends to be important for risk comprehension because it promotes more thorough, complex, and durable information representations, improving accuracy. Oudhoff and Timmermans (2015) found similar results in medical students with low and high numeracy. In particular, the authors showed that static icon arrays increased the amount of time that these participants invested when answering questions about perceived likelihoods of risk. The authors, however, did not investigate whether response times influence accuracy of risk understanding (see also Okan, Galesic, & Garcia-Retamero, 2016, for similar results in people with low and high graph literacy).

By influencing deliberation and elaborative encoding, visual aids can also support skilled judgment self-assessment and calibration. That is, visual aids can support and improve precise cognitive self-evaluation, thereby reducing overconfidence and improving various aspects of cognitive monitoring. A study conducted by Garcia-Retamero, Cokely, and Hoffrage (2015) showed that this improvement, in turn, increases accuracy of risk understanding. Participants in the study were a large sample of patients in Spain who made diagnostic inferences about medical tests on the basis of information about prevalence of a disease and the sensitivity and false-positive rate of the tests. For instance, patients had to infer the probability that a woman had breast cancer given that she had a positive mammography. Patients received information about the prevalence of breast cancer in women in the population and the sensitivity and false-positive rate of the mammography. As in the studies described earlier, half of the patients in this study received the information in numbers without a visual aid, whereas the other half received visual aids representing the numerical information. These patients received a grid like the one depicted in Figure 2.

When information about the prevalence of the disease and the medical tests was presented numerically, patients with high numeracy more appropriately assessed the accuracy of their diagnostic inferences (i.e., they had a better sense of how well they understood the information—they were less biased). That is, patients with high numeracy showed less extreme confidence in their comprehension and did not stop deliberating until they actually understood the diagnostic information, which in part explained why they made more accurate inferences than patients with low numeracy. However, when all patients received the visual aid representing the numerical information, most patients avoided overconfidence and made better inferences, including those with low numeracy. These results indicate that visual aids helped patients evaluate diagnostic information by helping them evaluate their own understanding of the relevant risks. Garcia-Retamero, Cokely, Ghazal, et al. (2016) showed similar results in people with low and high graph literacy.

Beyond general deliberation and accurate self-assessment, individual differences in numeracy and graph literacy also predict more skilled use of adaptive strategies for exploration and evaluation of graphical information. For instance, compared with people with low graph literacy, people with high graph literacy are more strategic when allocating attention to regions of the visual aids containing essential information for accurate interpretation (e.g., they attend more to titles, labels, and scales; Okan et al., 2016). As a result, it is more likely that people with high graph literacy evaluate essential information and extract more complex or precise knowledge from visual aids (e.g., it is more likely that they identify the referents of the concepts depicted; Okan et al., 2015; Woller-Carter et al., 2012; see also Carpenter & Shah, 1998; Shah & Freedman, 2011). This partially explains why people with at least a minimum level of graph literacy tend to profit more from the use of visual aids. People with high graph literacy also show lower reliance on irrelevant spatial features when interpreting visual aids, which also explains differences in accuracy of graph understanding between these people and those with low graph literacy (Okan, Garcia-Retamero, Galesic, & Cokely, 2012). Hess, Visschers, Siegrist, et al. (2011) showed similar results in people with low and high numeracy.

Greater deliberation, monitoring, and strategic exploration tend to work in concert to foster more integrated and representative risk understanding. Broadly, it may be useful to generally assume that more numerate people tend to process graphical (and risk-relevant) information more precisely and thoroughly. With evidence supporting this hypothesis, Hess, Visschers, and Siegrist (2011) showed that people with high numeracy even tend to count icons in icon arrays and can draw relevant meaning from this process (i.e., more precise comparison based on exact numbers). In contrast, people with low numeracy process visual aids holistically and tend to get confused when they are guided toward counting icons. To the extent that risk understanding is essential to decision making or health outcomes, these findings indicate that presenting visual aids to less numerate people should tend to be most beneficial when gross elements of the information need to be compared and considered (Feldman-Stewart et al., 2000; Zikmund-Fisher et al., 2010) or when recognition of superordinate classes and evaluation of part-to-whole relations in the data are informative (Reyna & Brainerd, 2008). In these ways and others, by providing visual explanatory cues and contextual information, visual aids should tend to enhance the impact of the numerical information while preserving simplicity (Hildon et al., 2012; Okan et al., 2015; Zikmund-Fisher et al., 2010).

Because visual aids cause relatively robust changes in risk understanding by shaping and fine-tuning knowledge representation in long-term memory, visual aids also tend to give rise to more enduring changes in attitudes and behavioral intentions, which can directly affect decision making and healthy behavior (see Figure 6; see also Garcia-Retamero & Dhami, 2013). The studies conducted by Garcia-Retamero and Cokely (2011, 2014a, 2015b) described earlier showed that visual aids can promote healthy behavior (prevention and illness-detecting behavior related to sexually transmitted diseases). This research also revealed that the increased condom use and adherence to screening was a result of changes in attitudes toward the behaviors and intentions to perform these behaviors. Ultimately, attitudes and behavioral intentions affected the likelihood of performing the healthy behaviors (see Garcia-Retamero & Cokely, 2015a, for a review).

Ongoing research conducted in large samples of patients with low numeracy further indicates that visual aids that improve risk understanding may also increase information trust and, as a result, can increase patients' willingness to participate in shared decision making (Garcia-Retamero & Cokely, 2016; Petrova et al., 2015). This benefit of visual aids appears to be particularly influential for vulnerable people with low numeracy, who generally tend to be passive in health decision making (Galesic & Garcia-Retamero, 2011b). Taken together, the conceptual framework depicted in Figure 6 shows how skills and visual aids work together to promote a more precise and accurate understanding of relevant risks, thereby influencing attitudes, behavioral intentions, and trust, which lead to healthier decisions and more positive health outcomes.

GRAPH PREFERENCES, OPINIONS ABOUT VISUAL AIDS, AND INDIVIDUAL DIFFERENCES IN SKILLS

Authors of several studies have investigated whether people with different levels of numeracy and/or graph literacy differ in their subjective perceptions about visual aids. In this section, we summarize this work and investigate whether results match those in objective performance. Studies including subjective measures have typically framed research questions very differently. Therefore, for the sake of consistency, in this section we limit our focus to studies involving graph preferences and/or opinions about accessibility of information, attractiveness of information, and information trust (i.e., 73% of the studies including subjective measures). In 81% and 38% of these studies, respectively, authors investigated and reported preferences/opinions in people with different levels of numeracy and graph literacy. The rest of the studies did not investigate whether levels of numeracy and/or graph literacy affected people's preferences or opinions, or authors did not report results. In 69% of these studies, authors investigated both preferences/opinions and objective performance (i.e., accuracy of risk understanding or accuracy of risk perceptions).

Results across these studies tend to be contradictory. Some studies (e.g., Gaissmaier et al., 2012; Oudhoff & Timmermans, 2015; Tait et al., 2010a) showed that visual aids are perceived as more attractive, helpful, effective, and trustworthy than numerical information. Yet other studies (e.g., Ruiz et al., 2013; Tait et al., 2012; Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin, 2010b) showed few or no discrepancies in people's opinions when they received visual aids as compared with numerical information reported in text. Broadly, people tend to prefer simple, static visual aids that do not depict unnecessary information (Zikmund-Fisher et al., 2010, 2012; Zikmund-Fisher, Witteman, et al., 2014; but see Okan et al., 2015; Tait et al., 2012). Compared with other visual aids, people had the least experience with icon arrays and felt that they preferred them the least (Hamstra et al., 2015)—a result that seems to be inconsistent with results for the objective measures summarized earlier. In addition, although individuals who had low numeracy and/or high graph literacy generally benefit from the use of visual aids, authors of only 23% (17%) of the studies concluded that these individuals evaluated visual aids more positively than numbers. In contrast, 54% (67%) of the studies showed few or no discrepancies between participants with low and high numeracy (or graph literacy) in preferences and/or opinions about visual aids.

Finally, results in 87% of the studies showed that information formats that contributed to enhance accuracy of performance were not always preferred or evaluated more positively than those that did not, and vice versa. That is, people's opinions and preferences do not robustly translate into actual improvements in risk understanding in the available literature; 38% of the studies showed that this was especially the case in participants with low numeracy and/or low graph literacy, who often showed larger discrepancies between results in subjective and objective measures; another 38% of the studies showed similar discrepancies in people with low and high numeracy/graph literacy; and the rest of the studies showed larger discrepancies in people with high numeracy/graph literacy (8%) or did not report results (15%; i.e., discrepancies could not be computed or inferred). We elaborate on the implications of these results in the following section.

EFFICIENT DESIGN OF VISUAL AIDS

In this article we review a large body of converging evidence from several multidisciplinary scholars showing that many people have problems understanding simple, health-relevant numerical expressions of probability about health. This research also shows that people who have well-developed numerical skills tend to have more accurate perceptions of health-related benefits and risks, and often make better decisions that promote valuable health outcomes. The authors of the research reviewed in this paper critically examined one potential solution to this problem: Appropriately designed visual aids tend to promote risk literacy and healthy behavior, particularly among people with limited numerical skills who have at least a basic level of graph literacy (i.e., they understand basics of graph interpretation).

Given the extant data and statistical analyses, our conclusions are likely to be remarkably robust because they are based on a wide range of experimental studies conducted on diverse groups of people ($N = 27{,}885$) from 60 different countries, including practicing physicians, patients, general community samples, probabilistic national (i.e., representative) samples, online panels, highly educated young adults, and high-risk individuals (e.g., people with low numeracy and language proficiency, older adults with limited skills, and immigrants). These groups of people have different cultural backgrounds, education levels, languages, governments, and health systems, and yet across studies, the benefits of visual aids have been consistent and compelling. The results should also be useful as the studies reviewed in this article evaluated many different performance quality measures, including accuracy of perceptions of risk and risk reductions, risk avoidance, reduction of errors and biases, inferences about the predictive power of medical tests and treatment effectiveness, assessment of subjective confidence and judgment calibration (i.e., metacognition or thinking about thinking), evaluation and integration of emotions and trust, assessment and trajectories of health outcomes, preferences and memory of health information, and changes in attitudes, behavioral intentions, behaviors, and informed decision making.

Moreover, the essential findings have held across a wide range of conditions and manipulations. For instance, results held when visual aids were provided either instead of or in addition to numerical information and when using different types of visual aids as appropriate for different tasks (e.g., icon arrays, line plots, and bar charts, presenting either affected individuals only or these individuals and the entire population at risk). Results also held when visual aids were more or less abstract and when the numerical information was presented using different information formats (e.g., absolute or relative risks). Taken together, the evidence strongly suggests that appropriately designed visual aids generally offer a transparent, effective, and ethically desirable means of risk communication in ecological and naturalistic settings, improving both risk understanding and health-relevant decision making. Table 1 presents a summary of what we consider to be the best and most essential evidence-based heuristics for the design of transparent visual aids as presented in this paper.

Beyond the extensive integrative review (see Ancker et al., 2006; Garcia-Retamero & Cokely, 2013; Lipkus, 2007; Lipkus & Hollands, 1999; Spiegelhalter et al., 2011; Visschers, Meertens, Passchier, & De Vries, 2009, for other examples), the current manuscript adds to the growing body of research on the efficacy of visual aids in three other noteworthy ways. First, our review shows that individual differences in basic numerical and graphical representation are essential when assessing the effectiveness of interventions including visual aids. Efficient design of decision support technologies requires an interacting systems perspective for efficient custom-tailored risk communication design. That is, good design requires careful attention to the fit between people (i.e., values, traits, responsibilities, skills), processes (e.g., decision strategies, heuristics), and their environments (e.g., information formats, affordances, cultural constraints). To the extent that visual aids leverage people's natural capacity to recognize and contextualize relationships in the data, visual aids are likely to confer significant judgment and decision-making benefits, which should tend to translate into improved health outcomes. Similar benefits have been observed in interventions involving simple communication principles and improved presentation of health-relevant numerical information (e.g., presenting important information in the appropriate order; Cavanaugh et al., 2009; Peters et al., 2007); in interventions in people with low numeracy who evaluate risk communications involving small, evolutionary plausible groups of individuals (Garcia-Retamero & Galesic, 2011); and in interventions involving consequences of healthy behavior expressed in terms of life expectancy rather than as risk of disease (Galesic & Garcia-Retamero, 2011a).

Second, our review sheds light on the psychological mechanisms that give rise to the benefits of visual aids in people with limited numerical skills (i.e., the mechanisms mediating the effect of visual aids; see Figure 6). Research shows that visual aids can improve judgment and decision making and help promote healthy behavior by promoting deliberation, by improving self-assessment and reducing overconfidence, and by promoting strategic allocation of attention, which in turn fosters a more meaningful and representative understanding of health-relevant numerical information. By establishing

TABLE 1: Evidence-Based Heuristic Guidelines for the Design of Transparent Visual Aids

Characteristics of Transparent Visual Aids
Keep information simple and focused on essentials. • Keep visual aids simple (Gillan, Wickens, Hollands, & Carswell, 1998; Trevena et al., 2012). ○ Use clear captions and titles. ○ Avoid using shadows and truncated scales. ○ Use the same scale for comparison. • When depicting numerical information visually, "less is more" (Zikmund-Fisher, Fagerlin, & Ubel, 2010). ○ Present one message per visual aid. ○ Only important information should be represented visually. Focus on the two or three key messages that you would like to communicate, and depict them using different visual aids. ○ Use interactive or dynamic visual aids only if they are designed to promote encoding and elaborative processing (Okan, Garcia-Retamero, Cokely, & Maldonado, 2015). • Use clear language to describe visual information (Garcia-Retamero & Dhami, 2011). ○ Label visual aids with informative captions that include the depicted key message. ○ Translate the materials into the plain language familiar to the audience.
The best type of visual aid depends on the communication goal. • Use bar graphs to compare several data points (Fischhoff, Brewer, & Downs, 2012; Lipkus, 2007; Lipkus & Hollands, 1999). • Use line graphs to depict trends over time (Fischhoff et al., 2012; Lipkus, 2007; Lipkus & Hollands, 1999). • Use pie graphs to communicate information about proportions (Fischhoff et al., 2012). • Use grids to depict very large numbers (Garcia-Retamero & Hoffrage, 2013; Garcia-Retamero, Cokely, & Hoffrage, 2015). • Use magnifier risk scales (including a magnifying lens) to depict very small numbers (Ancker, Senathirajah, Kukafka, & Starren, 2006). • Use icon arrays to communicate treatment risk reduction or risk of side effects. ○ Represent results of the baseline risk (i.e., sample of nontreated individuals) and the incremental/reduced risk due to treatment (i.e., sample of treated individuals) using different icon arrays (Galesic, Garcia-Retamero, & Gigerenzer, 2009; Hamstra et al., 2015; Hawley et al., 2008). ○ Depict both affected individuals (numerators) and the entire population at risk (denominators; Garcia-Retamero & Galesic, 2010b; Stone et al., 2003; Stone, Gabard, Groves, & Lipkus, 2015). ○ Keep the size of denominators (i.e., entire population at risk) in the treated and nontreated groups of individuals constant for comparison (Garcia-Retamero & Galesic, 2009; Garcia-Retamero, Galesic, & Gigerenzer, 2010; Okan, Garcia-Retamero, Cokely, & Maldonado, 2012). ○ Arrange icon arrays as groups in a block rather than in a random scattering (Ancker, Weber, & Kukafka, 2011a; Feldman-Stewart, Kocovski, McConnell, Brundage, & Mackillop, 2000; Zikmund-Fisher et al., 2012). ○ Use personlike icon arrays if possible (Zikmund-Fisher, Witteman, et al., 2014). Otherwise, icon arrays with more abstract icons might also be helpful (Gaissmaier et al., 2012).

(continued)

TABLE 1: (continued)

Depict numerical information in addition to visual aids.

- Add important numerical and textual information describing the information depicted in the visual aids to improve accuracy (Hamstra et al., 2015; Okan et al., 2015).
- Pay attention to the frame of the information.
  - ○ Use absolute rather than relative risks (Garcia-Retamero & Galesic, 2010b).
  - ○ Use frequencies rather than decimals, fractions, or percentages (Hoffrage, Gigerenzer, Krauss, & Martignon, 2002).
  - ○ Use both positive and negative frames (Garcia-Retamero & Galesic, 2010a; Garcia-Retamero & Cokely, 2011, 2014a, 2015b).
  - ○ Keep time frames constant (Fischhoff et al., 2012).

Effective communications anticipate user needs and skills.

- Learn about your target group and use the appropriate reading level (Garcia-Retamero & Dhami, 2011).
- Visual aids are especially useful for people with relatively low numeracy, those with limited knowledge about medical facts, and people with low levels of education (Garcia-Retamero & Cokely, 2014a; Garcia-Retamero & Galesic, 2009, 2010b; Garcia-Retamero, Cokely, Wicki, & Joeris, 2016).
  - ○ Objective and subjective numeracy scales can help quickly identify individuals with low and high numeracy (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012; Fagerlin, Zikmund-Fisher, et al., 2007; Lipkus, Samsa, & Rimer, 2001; Schapira et al., 2012; Schwartz, Woloshin, Black, & Welch, 1997; Weller et al., 2013).
- Even among people with high numeracy, there are variations in the ability to extract data and meaning from visual aids (i.e., graph literacy; Gaissmaier et al., 2012; Okan et al., 2015; Okan, Galesic, & Garcia-Retamero, 2016; Okan, Garcia-Retamero, Cokely, et al., 2012; Zikmund-Fisher, Witteman, et al., 2014).
  - ○ Objective and subjective graph literacy scales can help quickly identify individuals with low and high graph literacy (Galesic & Garcia-Retamero, 2011c; Garcia-Retamero, Cokely, Ghazal, & Joeris, 2016b; Garcia-Retamero, Petrova, Feltz, & Cokely, in press).
  - ○ To improve risk understanding in people with low graph literacy . . .
    - Use simple labels and explanations to convey the meaning of important information in the visual aid.
    - Encourage active, elaborative information processing by including reflective questions about the visual information, followed by accuracy feedback whenever practical (Okan et al., 2015).
  - ○ If this strategy is not feasible . . .
    - Use visual aids only with people who have a moderate-to-high level of graph literacy. People who lack basic graph literacy may be better off with numbers (Ruiz et al., 2013).

Scale validation studies improve high-stakes interventions.

- Validate visual aids before conducting an intervention by soliciting feedback and conducting usability studies including cognitive process tracing (e.g., eye tracking, verbal protocols; Okan et al., 2016; Trevena et al., 2012; Woller-Carter, Okan, Cokely, & Garcia-Retamero, 2012).
- Involve the target audience in the design, evaluation, and dissemination of visual aids (Hesse & Shneiderman, 2007).

enduring attitudes and fostering behavioral intentions, and by increasing trust, visual aids may further promote understanding and self-assessment. Research also shows that there is a tight causal link between individual differences in skills, self-regulation, metacognition, encoding, risk understanding, and decision making (Cokely & Kelley, 2009; Ghazal et al., 2014). More generally, the reviewed research extends work documenting the beneficial effects of designing information formats to promote active learning (Ancker et al., 2011a; Natter & Berry, 2005).

Third, our review shows that people's opinions and preferences do not typically translate into actual improvements in risk understanding. This conclusion accords with previous research documenting mismatches between evaluations of visual aids and objective performance (see Feldman-Stewart et al., 2000; McCaffery et al., 2012; Micallef, Dragicevic, & Fekete, 2012). Research suggests that it is not advisable to rely solely on people's preferences and opinions about visual aids when designing risk communications regardless of their skills. People may be more likely to use decision aids if graphical formats are customized to their preferences and/or opinions (Ancker et al., 2006). However, laypeople are not consistently able to select the information format that helps them improve accuracy according to their level of numeracy and/or graph literacy—a result that underscores the value in designing communications that are effective and well liked. Therefore, we recommend that practitioners and scientists collaborate on validation studies before conducting interventions. To enhance generalizability, we recommend representative rather than convenience sampling, whenever practically feasible. Validation studies offer the added benefit of user feedback and usability assessment, which may also benefit by drawing on assessments of individual differences (e.g., fast numeracy tests, like those at riskliteracy.org) and cognitive process tracing studies (e.g., eye tracking, verbal protocols; Okan et al., 2016; Trevena et al., 2012; Woller-Carter et al., 2012).

Although visual aids are increasingly being used and recommended for the communication of risks to the public (Garcia-Retamero & Cokely, 2013), some caution is warranted because the use of poorly designed or distorted graphs can be used to bias decision makers (Okan et al., 2016; Okan, Garcia-Retamero, Galesic, et al., 2012; Stone et al., 2015; Woller-Carter et al., 2012). Perhaps the most prominent potential concern when using visual aids is that graphs can include misleading features (Beattie & Jones, 1992, 2002; Kosslyn, 2006). To illustrate, Cooper, Schriger, Wallace, Mikulich, and Wilkes (2003) investigated design features of pharmaceutical advertisement graphs presented in medical journals. In a sample of 74 graphs, they found that almost 40% included distortions that confused rather than helped readers (see also Okan et al., 2016). The most common distortions were misuse of 3-D objects (20%), improper scale or improperly split axes (16%), and incorrect baselines (12%). Improperly designed graphs can lead to interpretive errors that, in turn, can affect important financial, medical, and legal decisions (Arunachalam, Pei, & Steinbart, 2002; Cooper et al., 2003).

To consistently and efficiently design visual aids, and to avoid potentially misleading graphs, designers should (a) follow principles of good graph design (e.g., Edward, 2001; Gillan, Wickens, Hollands, & Carswell, 1998; Kosslyn, 2006) and (b) consider the costs and benefits of validation studies conducted before interventions. At a minimum, designers should conduct brief heuristic evaluations to review proposed visual aids based on the guidelines provided in Table 1. Similar types of heuristic evaluations and reviews are industry standards in use by human factors engineers for related products, such as interactive visual displays in computer software, handheld electronics, and websites. Future design guidelines may include directions for creating graph interfaces specifically tailored to users (i.e., personalization of medical risk information based on one's level of graph literacy and risk literacy). Addressing these issues will require more interdisciplinary work at the intersections of human factors, cognitive science, decision science, and usability.

FRONTIERS AND ETHICAL POLICIES

Despite noteworthy progress, many fundamental and practical questions remain unanswered. The research that we reviewed here showed that visual aids can enhance risk comprehension but may fail to reach individuals disadvantaged by factors including limited graph literacy skills. A certain level of graph literacy seems to be necessary to associate the patterns contained in graphs with meaningful interpretations of the data depicted. Nevertheless, recent research indicates that visual aids generally offer a relatively efficient means of reaching diverse individuals regardless of their levels of numeracy and graph literacy, as long as the visual aids include simple labels and explanations that help convey meaning or are accompanied with training in the use of graphs (Okan et al., 2015). The available literature provides compelling evidence indicating that the substantial benefits of features supporting active, elaborative processing can extend to individuals across a very wide range of ability levels. This finding is consistent with the notion that the ability to understand and make good decisions about risks (i.e., risk literacy) does not generally require exceptional cognitive capacities but instead follows from the use of simple cognitive processes that can be effectively deployed by most people given sufficient support. Future studies should identify other suitable strategies for communicating health risks to people with low numeracy and low graph literacy. In this regard, ongoing work on the effect of analogies from people's lives shows that these information formats are undemanding in terms of numeracy and graph literacy, and are useful for promoting custom-tailored risk communication (Galesic & Garcia-Retamero, 2013).

On the basis of all relevant available data detailed in this review, we find that icon arrays tend to be the best "all purpose" type of visual aids. They may be particularly useful when it is not otherwise feasible to tailor the type of display to people's skills or the goal of the task. To maximize potential benefits of these and other visual aids, more research will be needed on other vulnerable groups (e.g., people in rural and inner-city areas, children, and people with mental illnesses) and other applications of visual aids (e.g., extreme weather conditions and transmission risk of deadly diseases). There is also a pressing need for more prospective studies on the comparative effects of visual aids in the long run (e.g., years after interventions). Authors of future research will also need to investigate age-related changes in numeracy and graph literacy throughout the life span and the implications of these changes for risk communication (Kleemans, Segers, & Verhoeven, 2011; Purpura, Hume, Sims, & Lonigan, 2011; Siegler & Booth, 2004; Siegler & Opfer, 2003). As well, there is an important opportunity to investigate the role of cultural factors in the development and expression of these skills and their influence on the efficacy of visual aids and general decision-making quality (Garcia-Retamero & Galesic, 2013).

As technology advances and as social and ecological climates continue to change, we will face new and unimagined risks. At the same time, we expect profound increases in access to information about personal and community risks, along with new opportunities to share in decision making and public policy processes. If we are serious about democratic and ethical ideals, like informed decision making, multidisciplinary teams will need to continue to push the frontier of efficient design and evaluation of risk communications. As documented in the current review, these kinds of investments can and often do result in major social and economic benefits, like improved health and well-being. We hope our review will support these and other related efforts, like the evolution of standards for the construction of decision aids (Fischhoff et al., 2012; Gillan et al., 1998; Trevena et al., 2012), the creation of free online tools for building better graphs (e.g., www.iconarray.com), and the mapping and modeling of risk literacy (e.g., "risk reading levels") and general decision-making skills around the world (e.g., www.riskliteracy.org). Sustainable promotion of informed decision making in the 21st century demands continuous innovation and improvement. To the extent the current findings generalize, we can expect such efforts to efficiently promote more inclusive access to life-altering information and enduring health.

APPENDIX

TABLE A1: Demographics of Participants

Studies Included in the Systematic Review	Sample Size	Type of Participants	Percentage of Males	Average Age	Percentage With High School or Less	Nationality
Garcia-Retamero, Cokely, Ghazal, & Joeris (2016)	309	Physicians and patients	57%	51	42%	48 countries
Garcia-Retamero, Cokely, Wicki, & Joeris (2016)	292	Physicians	82%	44	0%	60 countries
Okan, Galesic, & Garcia-Retamero (2016)	99	Young adults	44%	25	50%	Germany
Nayak et al. (2016)	50	Patients	100%	71	0%	United States
Chen & Yang (2015)	127	Young adults, highly educated	69%	21	0%	United States
Honda et al. (2015)	900	Online panel	0%	40	Not available	Japan
Garcia-Retamero, Cokely, & Hoffrage (2015b)	108	Patients	22%	52	86%	Spain
Stone, Gabard, Groves, & Lipkus (2015)	314	Young adults, highly educated	46%	Not available	0%	United States
Okan, Garcia-Retamero, Cokely, & Maldonado (2015)	458	Young adults, highly educated	33%	21	0%	Spain
Oudhoff & Timmermans (2015)	192	Young adults, highly educated	29%	22	0%	The Netherlands
Hamstra et al. (2015)	420	General community sample	100%	52	13%	United States
Zikmund-Fisher, Witteman, et al. (2014b)	1,502	Online panel	46%	54	23%	United States
Garcia-Retamero & Cokely (2014a)	646	High-risk individuals	46%	19	0%	Spain
Mason et al. (2014)	570	Online panel	49%	48	14%	United Kingdom
Ruiz et al. (2013)	120	Patients	100%	61	100%	United States
Bruine de Bruin, Stone, Gibson, Fischbeck, & Shoraka (2013)	234	Online panel	51%	39	Not available	United States
Tait, Voepel-Lewis, Brennan-Martinez, McGonegal, & Levine (2012)	200	Patients	47%	54	21%	United States

(continued)

TABLE A1: (continued)

Studies Included in the Systematic Review	Sample Size	Type of Participants	Percentage of Males	Average Age	Percentage With High School or Less	Nationality
Okan, Garcia-Retamero, Cokely, & Maldonado (2012)	168	Young adults, highly educated	84%	20	0%	Spain
Woller-Carter, Okan, Cokely, & Garcia-Retamero (2012)	73	Young adults, highly educated	64%	21	0%	United States
Zikmund-Fisher et al. (2012)	4,198	Online panel	54%	49	19%	United States
Gaissmaier et al. (2012)	275	Young adults, highly educated	47%	31	0%	Germany
Okan, Garcia-Retamero, Galesic, & Cokely (2012)	182	Online panel	46%	34	Not available	United States
Hess, Visschers, Siegrist, & Keller (2011)	47	General community sample	68%	52	51%	Switzerland
Zikmund-Fisher, Dickson, & Witteman (2011)	2,426	Online panel	50%	49	18%	United States
Brown et al. (2011)	120	Patients	0%	46	8%	United States
Hess, Visschers, & Siegrist (2011)	1,175	General community sample	47%	46	71%	Switzerland
Ancker, Weber, & Kukafka (2011b)	165	Online panel	36%	32	29%	United States
Zikmund-Fisher, Fagerlin, & Ubel (2010)	1,552	Online panel	0%	55	25%	United States
Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin (2010b)	408	Patients	26%	36	17%	United States
Garcia-Retamero & Galesic (2010a)	987	Probabilistic national samples	50%	48	76%	United States and Germany
Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin (2010a)	4,685	Online panel	43%	39	19%	United States
Garcia-Retamero & Galesic (2010b)	987	Probabilistic national samples	50%	48	76%	United States and Germany
Garcia-Retamero & Galesic (2009)	1,047	Probabilistic national samples	50%	49	76%	United States and Germany

(continued)

TABLE A1: (continued)

Studies Included in the Systematic Review	Sample Size	Type of Participants	Percentage of Males	Average Age	Percentage With High School or Less	Nationality
Keller & Siegrist (2009)	266	General community sample	0%	48	76%	Switzerland
Galesic, Garcia-Retamero, & Gigerenzer (2009)	171	Young adults, highly educated	46%	40	20%	Germany
Hawley et al. (2008)	2,412	Online panel	48%	49	18%	United States

Note. Studies are ordered by publication date. Studies using distorted graphs or visual aids depicting only number of people at risk were excluded for all the analyses in the tables (see text for a justification of this and the other criteria). One of the studies on perceptions of risks (Chen & Yang, 2015) did not include correlations between these perceptions and correct values or otherwise did not compare perceptions for different levels of probability. The study was excluded as it did not measure graph preferences or estimates about accessibility of information, attractiveness of information, or information trust.

TABLE A2: Basic Characteristics of the Reviewed Studies

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Garcia-Retamero, Cokely, Ghazal, & Joeris (2016)	Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero (2012)	Galesic & Garcia-Retamero (2011c); Garcia-Retamero, Cokely, Ghazal, et al. (2016)	Not investigated	Static icon arrays, with a control condition	Accuracy of risk understanding, metacognitive judgment calibration, and confidence	Investigated
Garcia-Retamero, Cokely, Wicki, & Joeris (2016)	Cokely et al. (2012)	Not measured	Not investigated	Static icon arrays, with a control condition	Accuracy of risk understanding, reading latency, and estimate latency	Investigated

(continued)

TABLE A2: (continued)

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Okan, Galesic, & Garcia-Retamero (2016)	Cokely et al. (2012); Lipkus, Samsa, & Rimer (2001); Schwartz, Woloshin, Black, & Welch (1997), not included in main analyses	Galesic & Garcia-Retamero (2011c)	Not investigated	Static bar and line graphs with conflicts (i.e., distorted graphs), without a control condition	Accuracy of risk understanding, and eye movements (allocation of attention on regions of visual aids containing essential information)	Investigated
Nayak et al. (2016)	Fagerlin, Zikmund-Fisher, et al. (2007)	Galesic & Garcia-Retamero (2011c)	Investigated	Static bars and lines graphs, without a control condition	Accuracy of risk understanding, and preferences of information format	Not investigated
Chen & Yang (2015)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Static bar graphs, with a control condition	Risk perceptions	Not investigated
Honda et al. (2015)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Static bar and line graph, with a control condition	Accuracy of risk understanding	Not investigated
Garcia-Retamero, Cokely, & Hoffrage (2015b)	Lipkus et al. (2001); Schwartz et al. (1997)	Not measured	Not investigated	Static visual grids, with a control condition	Accuracy of risk understanding, metacognitive judgment calibration, and self-assessments of accuracy of risk understanding	Investigated

(continued)

TABLE A2: (continued)

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Stone, Gabard, Groves, & Lipkus (2015)	Schwartz et al. (1997)	Not measured	Not investigated	Static icon arrays depicting affected individuals only (numerators), with a control condition	Accuracy of risk understanding, risk perceptions, and affective reactions	Not investigated
Okan, Garcia-Retamero, Cokely, & Maldonado (2015)	Cokely et al. (2012); Lipkus et al. (2001); Schwartz et al. (1997)	Galesic & Garcia-Retamero (2011c)	Not investigated	Dynamic icon arrays, without a control condition	Accuracy of risk understanding, accessibility of information format, preferences of information format, and confidence	Investigated
Oudhoff & Timmermans (2015)	Schwartz et al. (1997)	Not measured	Not investigated	Static bar graphs and icon arrays, with a control condition	Risk perceptions, decision making, estimate latency, and ease of imagining risk	Not investigated
Hamstra et al. (2015)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Different types of static visual aids (line and bar graphs, pies, icon arrays), without a control condition	Accuracy of risk understanding, preferences of information format, and familiarity	Not investigated
Zikmund-Fisher, Witteman, et al. (2014b)	Fagerlin, Zikmund-Fisher, et al. (2007)	Galesic & Garcia-Retamero (2011c)	Not investigated	Different types of static icon arrays that differ in iconicity, without a control condition	Risk perceptions, recall, accessibility of information format, and preferences of information format	Not investigated

(continued)

TABLE A2: (continued)

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Garcia-Retamero & Cokely (2014a)	Lipkus et al. (2001); Schwartz et al. (1997)	Galesic & Garcia-Retamero (2011c)	Investigated	Static icon arrays, with a control condition	Decision making, attractiveness of information, attitudes towards information, affective reactions towards information, and behavioral intentions	Investigated
Mason et al. (2014)	Lipkus et al. (2001); Schwartz et al. (1997)	Not measured	Not investigated	Interactive bar graph, with a control condition	Recall	Not investigated
Ruiz et al. (2013)	Lipkus et al. (2001); Schwartz et al. (1997)	Galesic & Garcia-Retamero (2011c)	Not investigated	Static icon arrays, with a control condition	Accuracy of risk understanding, recall, decision making, accessibility of information format, behavioral intentions, and confidence	Not investigated
Bruine de Bruin, Stone, Gibson, Fischbeck, & Shoraka (2013)	Lipkus et al. (2001); Schwartz et al. (1997)	Not measured	Not investigated	Static bar graphs with emphasis on denominators, with a control condition	Accuracy of risk understanding, decision making, attractiveness of information, affective reactions towards information, information trust, and behavioral intentions	Not investigated

(continued)

TABLE A2: (continued)

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Tait, Voepel-Lewis, Brennan-Martinez, McGonegal, & Levine (2012)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Different types of dynamic visual aids (bar graphs, pies, icon arrays), with a control condition	Accuracy of risk understanding, accessibility of information format, preferences of information format, and information trust	Not investigated
Okan, Garcia-Retamero, Cokely, & Maldonado (2012)	Not measured	Galesic & Garcia-Retamero (2011c)	Not investigated	Static icon arrays, without a control condition	Accuracy of risk understanding and confidence	Not investigated
Woller-Carter, Okan, Cokely, & Garcia-Retamero (2012)	Cokely et al. (2012)	Galesic & Garcia-Retamero (2011c)	Not investigated	Static bar graphs with conflicts (i.e., distorted graphs), without a control condition	Accuracy of risk understanding, recall, and eye movements (allocation of attention on regions of visual aids containing essential information)	Investigated
Zikmund-Fisher et al. (2012)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Dynamic icon arrays, without a control condition	Accuracy of risk understanding, decision making, accessibility of information format, and preferences of information format	Not investigated
Gaissmaier et al. (2012)	Lipkus et al. (2001); Schwartz et al. (1997), not included in main analyses	Galesic & Garcia-Retamero (2011c)	Not investigated	Different types of static icon arrays that differ in iconicity, with a control condition	Accuracy of risk understanding, recall, accessibility of information format, and attractiveness of information	Not investigated

(continued)

TABLE A2: (continued)

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Okan, Garcia-Retamero, Galesic, & Cokely (2012)	Lipkus et al. (2001); Schwartz et al. (1997), not included in main analyses	Galesic & Garcia-Retamero (2011c)	Not investigated	Static bar graphs with conflicts (i.e., distorted graphs), without a control condition	Accuracy of risk understanding	Investigated
Hess, Visschers, Siegrist, & Keller (2011)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Static graphical risk ladder (Paling perspective scale), without a control condition	Eye movements (allocation of attention on regions of visual aids containing essential information)	Not investigated
Zikmund-Fisher, Dickson, & Witteman (2011)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Interactive icon arrays, without a control condition	Accuracy of risk understanding and decision making	Not investigated
Brown et al. (2011)	Schwartz et al. (1997)	Not measured	Not investigated	Different types of static visual aids (vertical bar, horizontal bar, line, icon array), without a control condition	Accuracy of risk understanding and preferences of information format	Not investigated
Hess, Visschers, & Siegrist (2011)	Lipkus et al. (2001); a modified version of Fagerlin et al. (2007)	Not measured	Not investigated	Static icon arrays, without a control condition	Risk perceptions	Investigated
Ancker, Weber, & Kukafka (2011b)	A modified version of Lipkus et al. (2001)	Not measured	Not investigated	Interactive icon arrays, without a control condition	Risk perceptions, accessibility of information format, affective reactions, and behavioral intentions	Not investigated

(continued)

TABLE A2: (continued)

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Zikmund-Fisher, Fagerlin, & Ubel (2010)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Static icon arrays (depicting and not depicting unnecessary information), without a control condition	Accuracy of risk understanding, time to complete the task (cognitive effort), accessibility of information format, and preferences of information format	Not investigated
Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin (2010b)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Static icon arrays, with control condition	Accuracy of risk understanding and accessibility of information format	Not investigated
Garcia-Retamero & Galesic (2010a)	Lipkus et al. (2001); Schwartz et al. (1997)	Not measured	Not investigated	Different types of static visual aids (vertical bars, horizontal bars, pie charts, and icon arrays), with control condition	Risk perceptions	Not investigated
Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin (2010a)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Static icon arrays, with control condition	Accuracy of risk understanding, risk perceptions, accessibility of information format, information trust, and behavioral intentions	Not investigated

(continued)

TABLE A2: (continued)

Studies Included in the Systematic Review	Measure of Numeracy	Measure of Graph Literacy	Interaction Between Numeracy and Graph Literacy	Manipulation of Information Format	Dependent Variable(s)	Psychological Mechanisms
Garcia-Retamero & Galesic (2010b)	Lipkus et al. (2001); Schwartz et al. (1997)	Galesic & Garcia-Retamero (2011c)	Investigated	Static icon arrays and bar graphs depicting either the entire population at risk or affected individuals only, with control condition	Accuracy of risk understanding	Not investigated
Garcia-Retamero & Galesic (2009)	Lipkus et al. (2001); Schwartz et al. (1997)	Not measured	Not investigated	Static icon arrays, with control condition	Accuracy of risk understanding	Not investigated
Keller & Siegrist (2009)	Lipkus et al. (2001); Schwartz et al. (1997)	Not measured	Not investigated	Static icon arrays and Paling perspective scale, with control condition	Risk perceptions	Not investigated
Galesic, Garcia-Retamero, & Gigerenzer (2009)	Lipkus et al. (2001); Schwartz et al. (1997)	Not measured	Not investigated	Static icon arrays, with control condition	Accuracy of risk understanding	Not investigated
Hawley et al. (2008)	Fagerlin, Zikmund-Fisher, et al. (2007)	Not measured	Not investigated	Different types of static visual aids (icon arrays, pie, bar, sparkplug, and clock), without a control condition	Accuracy of risk understanding, decision making, accessibility of information format, and information trust	Not investigated

Note. Studies are ordered by publication date. Studies using distorted graphs or visual aids depicting only number of people at risk were excluded for all the analyses in the tables (see text for a justification of this and the other criteria). One of the studies on perceptions of risks (Chen & Yang, 2015) did not include correlations between these perceptions and correct values or otherwise did not compare perceptions for different levels of probability. The study was excluded as it did not measure graph preferences or estimates about accessibility of information, attractiveness of information, or information trust.

TABLE A3: Main Results of the Reviewed Studies

	Objective Performance				Subjective Perceptions		Match Between Objective Performance and Subjective Perceptions	
Studies Included in the Systematic Review	Visual Aids Tend to Be Effective[a]	Visual Aids Are More Effective for People With Low Numeracy[b]	People With Low Numeracy Achieve Lower Performance With Visual Aids[c]	Visual Aids Are More Effective for People With High Graph Literacy[b]	Results Differ in People With Low and High Numeracy[d]	Results Differ in People With Low and High Graph Literacy[d]	Overall[e]	Larger Discrepancies in People With Low Numeracy or Low Graph Literacy[f]
Garcia-Retamero, Cokely, Ghazal, & Joeris (2016)	Yes	Yes	Yes	Yes	Not included in analyses	Not included in analyses	Not included in analyses	Not included in analyses
Garcia-Retamero, Cokely, Wicki, & Joeris (2016)	Yes	Yes	No	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Okan, Galesic, & Garcia-Retamero (2016)	Not included in analyses	Not investigated	Not investigated	Not included in analyses	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Nayak et al. (2016)	Yes	Not investigated	Yes	Not investigated	Yes	No	No	No
Chen & Yang (2015)	Not investigated	Not investigated	Not investigated	Not investigated	Not included in analyses	Not investigated	Cannot be inferred	Cannot be inferred
Honda et al. (2015)	Yes	No	Yes	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Garcia-Retamero, Cokely, & Hoffrage (2015b)	Yes	Yes	No	Not investigated	Not included in analyses	Not investigated	Not included in analyses	Not included in analyses
Stone, Gabard, Groves, & Lipkus (2015)	Not included in analyses	Not included in analyses	Not included in analyses	Not investigated	Not included in analyses	Not investigated	Not included in analyses	Not included in analyses
Okan, Garcia-Retamero, Cokely, & Maldonado (2015)	Yes	Not investigated	Yes	Not investigated	No	Yes	No	No
Oudhoff & Timmermans (2015)	Yes	Yes	No	Not investigated	Not included in analyses	Not investigated	Not included in analyses	Not included in analyses
Hamstra et al. (2015)	Yes	Not investigated	Yes	Not investigated	No	Not investigated	No	Yes

(continued)

TABLE A3: (continued)

	Objective Performance				Subjective Perceptions		Match Between Objective Performance and Subjective Perceptions	
Studies Included in the Systematic Review	Visual Aids Tend to Be Effective[a]	Visual Aids Are More Effective for People With Low Numeracy[b]	People With Low Numeracy Achieve Lower Performance With Visual Aids[c]	Visual Aids Are More Effective for People With High Graph Literacy[b]	Results Differ in People With Low and High Numeracy[d]	Results Differ in People With Low and High Graph Literacy[d]	Overall[e]	Larger Discrepancies in People With Low Numeracy or Low Graph Literacy[f]
Zikmund-Fisher, Witteman, et al. (2014b)	Yes	Not investigated	Yes	Not investigated	No	No	No	Yes
Garcia-Retamero & Cokely (2014a)	Yes	Yes	Yes	Yes	Yes	Yes	Not included in analyses	Not included in analyses
Mason et al. (2014)	No	No	Yes	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Ruiz et al. (2013)	No	No	Yes	Yes	No	No	No	No
Bruine de Bruin, Stone, Gibson, Fischbeck, & Shoraka (2013)	No	No	Yes	Not investigated	Yes	Not investigated	No	No
Tait, Voepel-Lewis, Brennan-Martinez, McGonegal, & Levine (2012)	No	No	No	Not investigated	No	Not investigated	No	No
Okan, Garcia-Retamero, Cokely, & Maldonado (2012)	Yes	Not investigated	Not investigated	Not investigated	Not investigated	Not included in analyses	Not included in analyses	Not included in analyses
Woller-Carter, Okan, Cokely, & Garcia-Retamero (2012)	Not included in analyses	Not investigated	Not included in analyses	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred

(continued)

TABLE A3: (continued)

	Objective Performance				Subjective Perceptions		Match Between Objective Performance and Subjective Perceptions	
Studies Included in the Systematic Review	Visual Aids Tend to Be Effective[a]	Visual Aids Are More Effective for People With Low Numeracy[b]	People With Low Numeracy Achieve Lower Performance With Visual Aids[c]	Visual Aids Are More Effective for People With High Graph Literacy[b]	Results Differ in People With Low and High Numeracy[d]	Results Differ in People With Low and High Graph Literacy[d]	Overall[e]	Larger Discrepancies in People With Low Numeracy or Low Graph Literacy[f]
Zikmund-Fisher et al. (2012)	No	Not investigated	Yes	Not investigated	Yes	Not investigated	Yes	Only when result in previous column is yes
Gaissmaier et al. (2012)	Yes	Not investigated	Not investigated	Yes	Not investigated	Yes	No	No
Okan, Garcia-Retamero, Galesic, & Cokely (2012)	Not included in analyses	Not investigated	Not investigated	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Hess, Visschers, Siegrist, & Keller (2011)	Yes	Not investigated	No	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Zikmund-Fisher, Dickson, & Witteman (2011)	No	Not investigated	Yes	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Brown et al. (2011)	Yes	Not investigated	Yes	Not investigated	Not reported	Not investigated	No	Not reported
Hess, Visschers, & Siegrist (2011)	Yes	Not investigated	Yes	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Ancker, Weber, & Kukafka (2011b)	No	Not investigated	Yes	Not investigated	Yes	Not investigated	No	Yes
Zikmund-Fisher, Fagerlin, & Ubel (2010)	Yes	Not investigated	Yes	Not investigated	Yes	Not investigated	No	Yes

(continued)

TABLE A3: (continued)

Studies Included in the Systematic Review	Objective Performance				Subjective Perceptions		Match Between Objective Performance and Subjective Perceptions	
	Visual Aids Tend to Be Effective[a]	Visual Aids Are More Effective for People With Low Numeracy[b]	People With Low Numeracy Achieve Lower Performance With Visual Aids[c]	Visual Aids Are More Effective for People With High Graph Literacy[b]	Results Differ in People With Low and High Numeracy[d]	Results Differ in People With Low and High Graph Literacy[d]	Overall[e]	Larger Discrepancies in People With Low Numeracy or Low Graph Literacy[f]
Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin (2010b)	Yes	No	Yes	Not investigated	No	Not investigated	No	No
Garcia-Retamero & Galesic (2010a)	Yes	Yes	Yes	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Tait, Voepel-Lewis, Zikmund-Fisher, & Fagerlin (2010a)	Yes	Yes	Yes	Not investigated	Not reported	Not investigated	Yes	Only when result in previous column is yes
Garcia-Retamero & Galesic (2010b)	Yes	Yes	Yes	Yes	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Garcia-Retamero & Galesic (2009)	Yes	Yes	Yes	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Keller & Siegrist (2009)	No	No	Yes	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred

(continued)

TABLE A3: (continued)

Studies Included in the Systematic Review	Objective Performance				Subjective Perceptions		Match Between Objective Performance and Subjective Perceptions	
	Visual Aids Tend to Be Effective[a]	Visual Aids Are More Effective for People With Low Numeracy[b]	People With Low Numeracy Achieve Lower Performance With Visual Aids[c]	Visual Aids Are More Effective for People With High Graph Literacy[b]	Results Differ in People With Low and High Numeracy[d]	Results Differ in People With Low and High Graph Literacy[d]	Overall[e]	Larger Discrepancies in People With Low Numeracy or Low Graph Literacy[f]
Galesic, Garcia-Retamero, & Gigerenzer (2009)	Yes	Yes	No	Not investigated	Not investigated	Not investigated	Cannot be inferred	Cannot be inferred
Hawley et al. (2008)	Yes	Not investigated	Yes	Not investigated	Yes	Not investigated	No	No

Note. Studies are ordered by publication date. Studies using distorted graphs or visual aids depicting only number of people at risk were excluded for all the analyses in the tables (see text for a justification of this and the other criteria). One of the studies on perceptions of risks (Chen & Yang, 2015) did not include correlations between these perceptions and correct values or otherwise did not compare perceptions for different levels of probability. The study was excluded as it did not measure graph preferences or estimates about accessibility of information, attractiveness of information, or information trust.

[a]Yes if visual aids tend to improve performance regardless of participants' skills.

[b]Yes if visual aids are more effective for people with low versus high numeracy (or graph literacy) when comparing the visual aids with numbers in text. Studies that did not involve the efficacy of visual aids as compared with presenting numerical information in written text (i.e., did not include a control condition) were excluded.

[c]Yes if people with low numeracy achieved lower performance than those with high numeracy when receiving visual aids.

[d]Yes if people with low versus high numeracy (or graph literacy) differ in their graph preferences and/or subjective opinions about visual aids. Only studies measuring graph preferences and/or estimates about accessibility of information, attractiveness of information, and information trust were included.

[e]Yes if results in subjective perceptions match those in objective performance (i.e., if participants prefer or have a more positive opinion about the information format(s) that help them be more accurate). Only studies including the following objective and subjective measures were considered: (a) accuracy of risk understanding and/or risk perceptions (correlated with objective measures or compared when using different probability levels; objective measures) and (b) graph preferences and/or estimates about accessibility of information, attractiveness of information, and information trust (subjective measures).

[f]Yes if results in subjective perceptions do not match those in objective performance and participants with low numeracy (or low graph literacy) showed larger discrepancies between objective performance and subjective perceptions than participants with low numeracy (or graph literacy). Only studies including the objective and subjective measures mentioned in the previous note were considered.

ACKNOWLEDGMENTS

The current research was funded by the Ministerio de Economía y Competitividad (Spain) (PSI2011-22954 and PSI2014-51842-R), the Ministerio de Educación y Ciencia (Spain) (PSI2008-02019), the Foundation for Informed Medical Decision Making (United States), the National Science Foundation (United States) (SES-1253263), and the Max Planck Society (Germany). This work was also supported by the Harvard Center for Risk Analysis, Boston, Massachusetts. The authors declare independence from these funding agencies and do not have conflicts of interest including financial interests, activities, relationships, and affiliations.

KEY POINTS

- A conceptual framework shows that numeracy affects health outcomes by improving risk understanding and promoting shared decision making.
- A systematic review shows that visual aids are often remarkably beneficial for diverse people with different levels of numeracy and graph literacy.
- Well-designed visual aids robustly improve risk understanding by encouraging more thorough deliberation, facilitating self-assessment, and reducing biased risk representations, which in turn benefit attitudes, behavioral intentions, and trust, leading to healthier decisions and more positive health outcomes.
- Five essential categories of evidence-based guidelines for heuristic evaluation and design of effective visual aids are identified.
- Conclusions are based on large and diverse groups of people ($N = 27{,}885$) from 60 different countries, including practicing physicians, patients, general community samples, probabilistic national (i.e., representative) samples, online panels, highly educated young adults, and high-risk individuals.

REFERENCES

References marked with an asterisk indicate studies included in the systematic review.

Abdel-Kader, K., Dew, M. A., Bhatnagar, M., Argyropoulos, C., Karpov, I., Switzer, G., & Unruh, M. L. (2010). Numeracy skills in CKD: Correlates and outcomes. *Clinical Journal of the American Society of Nephrology*, *5*, 1566–1573.

Ancker, J. S., & Kaufman, D. (2007). Rethinking health numeracy: A multidisciplinary literature review. *Journal of the American Medical Informatics Association*, *14*, 713–721.

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.

Ancker, J. S., Weber, E. U., & Kukafka, R. (2011a). Effect of arrangement of stick figures on estimates of proportion in risk graphics. *Medical Decision Making*, *31*, 143–150.

*Ancker, J. S., Weber, E. U., & Kukafka, R. (2011b). Effects of game-like interactive graphics on risk perceptions and decisions. *Medical Decision Making*, *31*, 130–142.

Anderson, B. L., Obrecht, N. A., Chapman, G. B., Driscoll, D. A., & Schulkin, J. (2011). Physicians' communication of Down syndrome screening test results: The influence of physician numeracy. *Genetics in Medicine*, *13*, 744–749.

Anderson, B. L., & Schulkin, J. (2014a). *Numerical reasoning in judgments and decision making about health*. Cambridge, UK: Cambridge University Press.

Anderson, B. L., & Schulkin, J. (2014b). Physicians' understanding and use of numeric information. In B. L. Anderson & J. Schulkin (Eds.), *Numerical reasoning in judgments and decision making about health* (pp. 59–79). Cambridge, UK: Cambridge University Press.

Apter, A. J., Cheng, J., Small, D., Bennett, I. M., Albert, C., Fein, D. G., . . . Van Horne, S. (2006). Asthma numeracy skill and health literacy. *Journal of Asthma*, *43*, 705–710.

Apter, A. J., Wang, X., Bogen, D., Bennett, I. M., Jennings, R. M., Garcia, L., . . . Ten Have, T. (2009). Linking numeracy and asthma-related quality of life. *Patient Education and Counseling*, *75*, 386–391.

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

Beattie, V., & Jones, M. J. (1992). The use and abuse of graphs in annual reports: Theoretical framework and empirical study. *Accounting and Business Research*, *22*, 291–303.

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

Brewer, N. T., Chapman, G. B., Gibbons, F. X., Gerrard, M., McCaul, K. D., & Weinstein, N. D. (2007). Meta-analysis of the relationship between risk perception and health behavior: The example of vaccination. *Health Psychology*, *26*, 136–145.

*Brown, S. M., Culver, J. O., Osann, K. E., MacDonald, D. J., Sand, S., Thornton, A. A., . . . Burke, H. B. (2011). Health literacy, numeracy, and interpretation of graphical breast cancer risk estimates. *Patient Education and Counseling*, *83*, 92–98.

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.

*Bruine de Bruin, W., Stone, E. R., Gibson, J. M., Fischbeck, P. S., & Shoraka, M. B. (2013). The effect of communication design and recipients' numeracy on responses to UXO risk. *Journal of Risk Research*, *16*, 981–1004.

Brundage, M., Feldman-Stewart, D., Leis, A., Bezjak, A., Degner, L., Velji, K., . . . Pater, J. (2005). Communicating quality of life information to cancer patients: A study of six presentation formats. *Journal of Clinical Oncology*, *23*, 6949–6956.

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.

Cavanaugh, K., Huizinga, M. M., Wallston, K. A., Gebretsadik, T., Shintani, A., Davis, D., . . . Cherrington, A. (2008). Association of numeracy and diabetes control. *Annals of Internal Medicine*, *148*, 737–746.

Cavanaugh, K., Wallston, K. A., Gebretsadik, T., Shintani, A., Huizinga, M. M., Davis, D., . . . Rothman, R. L. (2009). Addressing literacy and numeracy to improve diabetes care: Two randomized controlled trials. *Diabetes Care*, *32*, 2149–2155.

*Chen, Y., & Yang, Z. J. (2015). Message formats, numeracy, risk perceptions of alcohol-attributable cancer, and intentions for binge drinking among college students. *Journal of Drug Education*, *45*, 37–55.

Cokely, E. T., Feltz, A., Ghazal, S., Allan, J. N., Petrova, D., & Garcia-Retamero, R. (in press). Decision making skill: From intelligence to numeracy and expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *Cambridge handbook of expertise and expert performance* (2nd ed.). New York, NY: Cambridge University Press.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, *7*, 25–47.

Cokely, E. T., Ghazal, S., & Garcia-Retamero, R. (2014). Measuring numeracy. In B. L. Anderson & J. Schulkin (Eds.), *Numerical reasoning in judgments and decision making about health* (pp. 11–38). Cambridge, UK: Cambridge University Press.

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. L. (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.

Cox, D. S., Cox, A. D., Sturm, L., & Zimet, G. (2010). Behavioral interventions to increase HPV vaccination acceptability among mothers of young girls. *Health Psychology*, *29*, 29–39.

Davids, S. L., Schapira, M. M., McAuliffe, T. L., & Nattinger, A. B. (2004). Predictors of pessimistic breast cancer risk perceptions in a primary care population. *Journal of General Internal Medicine*, *19*, 310–315.

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

Dowse, R., & Ehlers, M. (2005). Medicine labels incorporating pictograms: Do they influence understanding and adherence? *Patient Education and Counseling*, *58*, 63–70.

Edward, T. (2001). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.

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

Estrada, C. A., Barnes, V., Collins, C., & Byrd, J. C. (1999). Health literacy and numeracy. *Journal of the American Medical Association*, *282*, 527–527.

Estrada, C. A., Martin-Hryniewicz, M., Peek, B. T., Collins, C., & Byrd, J. C. (2004). Literacy and numeracy skills and anticoagulation control. *American Journal of the Medical Sciences*, *328*, 88–93.

Fagerlin, A., Ubel, P. A., Smith, D. M., & Zikmund-Fisher, B. J. (2007). Making numbers matter: Present and future research in risk communication. *American Journal of Health Behavior*, *31*(Suppl. 1), S47–S56.

Fagerlin, A., Wang, C., & Ubel, P. A. (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.

Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the subjective numeracy scale. *Medical Decision Making*, *27*, 672–680.

Feldman-Stewart, D., Brundage, M. D., & Zotov, V. (2007). Further insight into the perception of quantitative information: Judgments of gist in treatment decisions. *Medical Decision Making*, *27*, 34–43.

Feldman-Stewart, D., Kocovski, N., McConnell, B. A., Brundage, M. D., & Mackillop, W. J. (2000). Perception of quantitative information for treatment decisions. *Medical Decision Making*, *20*, 228–238.

Fischhoff, B., Brewer, N. T., & Downs, J. S. (2012). *Communicating risks and benefits: An evidence based user's guide*. Washington, DC: Government Printing Office.

*Gaissmaier, W., Wegwarth, O., Skopec, D., Müller, A., Broschinski, S., & Politi, M. C. (2012). Numbers can be worth a thousand pictures: Individual differences in understanding graphical and numerical representations of health-related information. *Health Psychology*, *31*, 286–296.

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. (2011a). Communicating consequences of risky behaviors: Life expectancy versus risk of disease. *Patient Education and Counseling*, *82*, 30–35.

Galesic, M., & Garcia-Retamero, R. (2011b). Do low-numeracy people avoid shared decision making? *Health Psychology*, *30*, 336–341.

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

Galesic, M., & Garcia-Retamero, R. (2013). Using analogies to communicate information about health risks. *Applied Cognitive Psychology*, *27*, 33–42.

*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., Andrade, A., Sharit, J., & Ruiz, J. G. (2015). Is patient's numeracy related to physical and mental health? *Medical Decision Making*, *35*, 501–511.

Garcia-Retamero, R., & Cokely, E. T. (2011). 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*, *17*, 270–287.

Garcia-Retamero, R., & Cokely, E. T. (2013). Communicating health risks with visual aids. *Current Directions in Psychological Science*, *22*, 392–399.

*Garcia-Retamero, R., & Cokely, E. T. (2014a). The influence of skills, message frame, and visual aids on prevention of sexually transmitted diseases. *Journal of Behavioral Decision Making*, *27*, 179–189.

Garcia-Retamero, R., & Cokely, E. T. (2014b). Using visual aids to help people with low numeracy make better decisions. In B. L. Anderson & J. Schulkin (Eds.), *Numerical reasoning in judgments and decision making about health* (pp. 153–174). Cambridge, UK: Cambridge University Press.

Garcia-Retamero, R., & Cokely, E. T. (2015a). Brief messages to promote prevention and detection of sexually transmitted diseases. *Current HIV Research*, *13*, 408–420.

Garcia-Retamero, R., & Cokely, E. T. (2015b). Simple but powerful health messages for increasing condom use in young adults. *Journal of Sex Research*, *52*, 30–42.

Garcia-Retamero, R., & Cokely, E. T. (2016). *Promoting shared medical decision making in doctors and their patients*. Manuscript in preparation.

Garcia-Retamero, R., Cokely, E. T., & Ghazal, S. (2016). *Comparing risk literacy in 30 countries: New results from Berlin numeracy test*. Manuscript in preparation.

*Garcia-Retamero, R., Cokely, E. T., Ghazal, S., & Joeris, A. (2016). Measuring graph literacy without a test: A brief subjective assessment. *Medical Decision Making*, *36*, 854–867.

*Garcia-Retamero, R., Cokely, E. T., & Hoffrage, U. (2015). Visual aids improve diagnostic inferences and metacognitive judgment calibration. *Frontiers in Psychology*, *6*, 932.

*Garcia-Retamero, R., Cokely, E. T., Wicki, B., & Joeris, A. (2016). Improving surgeons' risk literacy with visual aids. *Patient Education and Counseling*, *99*, 1156–1161.

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*(s1), 46–57.

Garcia-Retamero, R., & Dhami, M. K. (2013). On avoiding framing effects in experienced decision makers. *Quarterly Journal of Experimental Psychology*, *66*, 829–842.

*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. (2013). *Transparent communication of health risks: Overcoming cultural differences*. New York, NY: Springer.

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

Garcia-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine*, *83*, 27–33.

Garcia-Retamero, R., & Muñoz, R. (2013). How to improve comprehension of medical risks in older adults. *Revista Latinoamericana De Psicología*, *45*, 253–264.

Garcia-Retamero, R., Petrova, D., Feltz, A., & Cokely, E. T. (in press). Measuring graph literacy: A systematic review. In R. Parrot (Ed.), *The Oxford encyclopedia of health and risk message design and processing*. Oxford, UK: Oxford University Press.

Garcia-Retamero, R., Wicki, B., Cokely, E. T., & Hanson, B. (2014). Factors predicting surgeons' preferred and actual roles in interactions with their patients. *Health Psychology*, *33*, 920–928.

Gardner, P. H., McMillan, B., Raynor, D. K., Woolf, E., & Knapp, P. (2011). The effect of numeracy on the comprehension of information about medicines in users of a patient information website. *Patient Education and Counseling*, *83*, 398–403.

Ghazal, S., Cokely, E. T., & Garcia-Retamero, R. (2014). Predicting biases in very highly educated samples: Numeracy and metacognition. *Judgment and Decision Making*, *9*, 15–34.

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.

Gillan, D. J., Wickens, C. D., Hollands, J. G., & Carswell, C. M. (1998). Guidelines for presenting quantitative data in HFES publications. *Human Factors*, *40*, 28–41.

Ginde, A. A., Clark, S., Goldstein, J. N., & Camargo, C. A., Jr. (2008). Demographic disparities in numeracy among emergency department patients: Evidence from two multicenter studies. *Patient Education and Counseling*, *72*, 350–356.

Goodyear-Smith, F., Arroll, B., Chan, L., Jackson, R., Wells, S., & Kenealy, T. (2008). Patients prefer pictures to numbers to express cardiovascular benefit from treatment. *Annals of Family Medicine*, *6*, 213–217.

Gurmankin, A. D., Baron, J., & Armstrong, K. (2004). Intended message versus message received in hypothetical physician risk communications: Exploring the gap. *Risk Analysis*, *24*, 1337–1347.

*Hamstra, D. A., Johnson, S. B., Daignault, S., Zikmund-Fisher, B. J., Taylor, J. M., Larkin, K., . . . Fagerlin, A. (2015). The impact of numeracy on verbatim knowledge of the longitudinal risk for prostate cancer recurrence following radiation therapy. *Medical Decision Making*, *35*, 27–36.

Hanoch, Y., Miron-Shatz, T., Cole, H., Himmelstein, M., & Federman, A. D. (2010). Choice, numeracy, and physicians-in-training performance: The case of Medicare part D. *Health Psychology*, *29*, 454–459.

*Hawley, S. T., Zikmund-Fisher, B., Ubel, P., Jancovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling*, *73*, 448–455.

*Hess, R., Visschers, V. H., & Siegrist, M. (2011a). Risk communication with pictographs: The role of numeracy and graph processing. *Judgment and Decision Making*, *6*, 263–274.

*Hess, R., Visschers, V. H., Siegrist, M., & Keller, C. (2011b). How do people perceive graphical risk communication? The role of subjective numeracy. *Journal of Risk Research*, *14*, 47–61.

Hesse, B. W., & Shneiderman, B. (2007). eHealth research from the user's perspective. *American Journal of Preventive Medicine*, *32*, S97–S103.

Hibbard, J. H., Peters, E., Dixon, A., & Tusler, M. (2007). Consumer competencies and the use of comparative quality information: It isn't just about literacy. *Medical Care Research and Review*, *64*, 379–394.

Higgins, J. P., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: Wiley.

Hildon, Z., Allwood, D., & Black, N. (2012). Impact of format and content of visual display of data on comprehension, choice and preference: A systematic review. *International Journal for Quality in Health Care*, *24*, 55–64.

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition*, *84*, 343–352.

*Honda, H., Ogawa, M., Murakoshi, T., Masuda, T., Utsumi, K., Park, S., . . . Wada, Y. (2015). Effect of visual aids and individual differences of cognitive traits in judgments on food safety. *Food Policy*, *55*, 33–40.

Huestegge, L., & Philipp, A. M. (2011). Effects of spatial compatibility on integration processes in graph comprehension. *Attention, Perception, & Psychophysics*, *73*, 1903–1915.

Huizinga, M. M., Beech, B. M., Cavanaugh, K. L., Elasy, T. A., & Rothman, R. L. (2008). Low numeracy skills are associated with higher BMI. *Obesity*, *16*, 1966–1968.

Kakkilaya, V., Groome, L. J., Platt, D., Kurepa, D., Pramanik, A., Caldito, G., . . . Davis, T. C. (2011). Use of a visual aid to improve counseling at the threshold of viability. *Pediatrics*, *128*, e1511–e1519.

Kaplan, R. M., & Frosch, D. L. (2005). Decision making in medicine and health care. *Annual Review of Clinical Psychology*, *1*, 525–556.

*Keller, C., & Siegrist, M. (2009). Effect of risk communication formats on risk perception depending on numeracy. *Medical Decision Making*, *29*, 483–490.

Kleemans, T., Segers, E., & Verhoeven, L. (2011). Cognitive and linguistic precursors to numeracy in kindergarten: Evidence from first and second language learners. *Learning and Individual Differences*, *21*, 555–561.

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

Kutner, M., Greenburg, E., Jin, Y., & Paulsen, C. (2006). *The health literacy of America's adults: Results from the 2003 national assessment of adult literacy* (NCES 2006-483). Washington, DC: National Center for Education Statistics.

Lindau, S. T., Basu, A., & Leitsch, S. A. (2006). Health literacy as a predictor of follow-up after an abnormal pap smear: A prospective study. *Journal of General Internal Medicine*, *21*, 829–834.

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. *Medical Decision Making*, *27*, 696–713.

Lipkus, I. M., & Hollands, J. G. (1999). The visual communication of risk. *Journal of the National Cancer Institute. Monographs*, *25*, 149–163.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, *21*, 37–44.

Lopez, K. D., Wilkie, D. J., Yao, Y., Sousa, V., Febretti, A., Stifter, J., . . . Keenan, G. M. (2016). Nurses' numeracy and graphical literacy: Informing studies of clinical decision support interfaces. *Journal of Nursing Care Quality*, *31*, 124–130.

Machtinger, E. L., Wang, F., Chen, L., Rodriguez, M., Wu, S., & Schillinger, D. (2007). A visual medication schedule to improve anticoagulation control: A randomized, controlled trial. *Joint Commission Journal on Quality and Patient Safety*, *33*, 625–635.

Manganello, J. A., & Clayman, M. L. (2011). The association of understanding of medical statistics with health information seeking and health provider interaction in a national sample of young adults. *Journal of Health Communication*, *16*(Suppl. 3), 163–176.

Martin, L. T., Haas, A., Schonlau, M., Derose, K. P., Rosenfeld, L., Rudd, R., & Buka, S. L. (2012). Which literacy skills are associated with smoking? *Journal of Epidemiology and Community Health*, *66*, 189–192.

*Mason, D., Boase, S., Marteau, T., Kinmonth, A. L., Dahm, T., Minorikawa, N., & Sutton, S. (2014). One-week recall of health risk information and individual differences in attention to bar charts. *Health, Risk & Society*, *16*, 136–153.

McCaffery, K. J., Dixon, A., Hayen, A., Jansen, J., Smith, S., & Simpson, J. M. (2012). The influence of graphic display format on the interpretations of quantitative risk information among adults with lower education and literacy: A randomized experimental study. *Medical Decision Making*, *32*, 532–544.

McCaffery, K. J., Smith, S. K., & Wolf, M. (2010). The challenge of shared decision making among patients with lower literacy: A framework for research and development. *Medical Decision Making*, *30*, 35–44.

Micallef, L., Dragicevic, P., & Fekete, J. (2012). Assessing the effect of visualizations on Bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics*, *18*, 2536–2545.

Moles, A. A. (1966). *Information theory and esthetic perception*. Urbana: University of Illinois Press.

Natter, H. M., & Berry, D. C. (2005). Effects of active information processing on the understanding of risk information. *Applied Cognitive Psychology*, *19*, 123–135.

*Nayak, J. G., Hartzler, A. L., Macleod, L. C., Izard, J. P., Dalkin, B. M., & Gore, J. L. (2016). Relevance of graph literacy in the development of patient-centered communication tools. *Patient Education and Counseling*, *99*, 448–454.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: Theory and practice. *Annals of Behavioral Medicine*, *35*, 261–274.

*Okan, Y., Galesic, M., & Garcia-Retamero, R. (2016). How people with low and high graph literacy process health graphs: Evidence from eye-tracking. *Journal of Behavioral Decision Making*, *29*, 271–294.

*Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2012). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making*, *25*, 390–401.

*Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2015). Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *Journal of Experimental Psychology: Applied*, *21*, 178–194.

*Okan, Y., Garcia-Retamero, R., Galesic, M., & Cokely, E. T. (2012). When higher bars are not larger quantities: On individual differences in the use of spatial information in graph comprehension. *Spatial Cognition & Computation*, *12*, 195–218.

Organisation for Economic Co-operation and Development. (2012). *PISA 2009 technical report*. Paris, France: OECD Publishing.

Osborn, C. Y., Cavanaugh, K., Wallston, K. A., & Rothman, R. L. (2010). Self-efficacy links health literacy and numeracy to glycemic control. *Journal of Health Communication*, *15*, 146–158.

*Oudhoff, J. P., & Timmermans, D. R. (2015). The effect of different graphical and numerical likelihood formats on perception of likelihood and choice. *Medical Decision Making*, *35*, 487–500.

Paasche-Orlow, M. K., & Wolf, M. S. (2007). The causal pathways linking health literacy to health outcomes. *American Journal of Health Behavior*, *31*(Suppl. 1), S19–S26.

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

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, *21*, 31–35.

Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology*, *108*, 802–822.

Peters, E., Dieckmann, N., Dixon, A., Hibbard, J. H., & Mertz, C. K. (2007). Less is more in presenting quality information to consumers. *Medical Care Research and Review*, *64*, 169–190.

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., Meilleur, L., & Tompkins, M. K. (2014). Numeracy and the Affordable Care Act: Opportunities and challenges. Appendix A. IOM (Institute of Medicine). In *Health literacy and* *numeracy: Workshop summary* (pp. 1–33). Washington, DC: National Academies Press.

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

Petrova, D., Garcia-Retamero, R., Catena, A., Cokely, E. T., Heredia, A., Arrebola, A., & Ramirez-Hernández, J. A. (in press). Numeracy predicts risk of pre-hospital decision delay: A retrospective study of acute coronary syndrome survival. *Annals of Behavioral Medicine*.

Petrova, D., Garcia-Retamero, R., Catena, A., & van der Pligt, J. (2016). To screen or not to screen: What factors influence complex screening decisions? *Journal of Experimental Psychology: Applied*, *22*, 247–260.

Petrova, D. G., Garcia-Retamero, R., & Cokely, E. T. (2015). Understanding harms and benefits of cancer screening: A model of factors that shape informed decision making. *Medical Decision Making*, *35*, 847–858.

Petrova, D. G., van der Pligt, J., & Garcia-Retamero, R. (2014). Feeling the numbers: On the interplay between risk, affect, and numeracy. *Journal of Behavioral Decision Making*, *27*, 191–199.

Portnoy, D. B., Roter, D., & Erby, L. H. (2010). The role of numeracy on client knowledge in BRCA genetic counseling. *Patient Education and Counseling*, *81*, 131–136.

Price, M., Cameron, R., & Butow, P. (2007). Communicating risk information: The influence of graphical display format on quantitative information perception: Accuracy, comprehension and preferences. *Patient Education and Counseling*, *69*, 121–128.

Purpura, D. J., Hume, L. E., Sims, D. M., & Lonigan, C. J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. *Journal of Experimental Child Psychology*, *110*, 647–658.

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

Reyna, V. F. (2008). A theory of medical decision making and health: Fuzzy trace theory. *Medical Decision Making*, *28*, 850–865.

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.

Rodríguez, V., Andrade, A. D., Garcia-Retamero, R., Anam, R., Rodríguez, R., Lisigurski, M., . . . Ruiz, J. G. (2013). Health literacy, numeracy, and graphical literacy among veterans in primary care and their effect on shared decision making and trust in physicians. *Journal of Health Communication*, *18*(Suppl. 1), 273–289.

Roter, D. L. (2005). Health literacy and the patient–provider relationship. In J. G. Schwartzberg, J. Van Geest, & C. C. Wang (Eds.), *Understanding health literacy: Implications for medicine and public health* (pp. 87–100). Chicago, IL: AMA Press.

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.

*Ruiz, J. G., Andrade, A. D., Garcia-Retamero, R., Anam, R., Rodriguez, R., & Sharit, J. (2013). Communicating global cardiovascular risk: Are icon arrays better than numerical estimates in improving understanding, recall and perception of risk? *Patient Education and Counseling*, *93*, 394–402.

Schapira, M. M., Walker, C. M., Cappaert, K. J., Ganschow, P. S., Fletcher, K. E., McGinley, E. L., . . . Jacobs, E. A. (2012). The numeracy understanding in medicine instrument: A measure of health numeracy developed using item response theory. *Medical Decision Making*, *32*, 851–865.

Schirillo, J. A., & Stone, E. R. (2005). The greater ability of graphical versus numerical displays to increase risk avoidance involves a common mechanism. *Risk Analysis*, *25*, 555–566.

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., & 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.

Sharit, J., Lisigurski, M., Andrade, A. D., Karanam, C., Nazi, K. M., Lewis, J. R., & Ruiz, J. G. (2014). The roles of health literacy, numeracy, and graph literacy on the usability of the VA's personal health record by veterans. *Journal of Usability Studies*, *9*, 173–193.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*, 237–243.

Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, *333*, 1393–1400.

*Stone, E. R., Gabard, A. R., Groves, A. E., & Lipkus, I. M. (2015). Effects of numerical versus foreground-only icon displays on understanding of risk magnitudes. *Journal of Health Communication*, *20*, 1230–1241.

Stone, E. R., Sieck, W. R., Bull, B. E., Frank Yates, J., 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.

Stone, E. R., Yates, J. F., & Parker, A. M. (1997). Effects of numerical and graphical displays on professed risk-taking behavior. *Journal of Experimental Psychology: Applied*, *3*, 243–256.

*Tait, A. R., Voepel-Lewis, T., Brennan-Martinez, C., McGonegal, M., & Levine, R. (2012). Using animated computer-generated text and graphics to depict the risks and benefits of medical treatment. *The American Journal of Medicine*, *125*, 1103–1110.

*Tait, A. R., Voepel-Lewis, T., Zikmund-Fisher, B. J., & Fagerlin, A. (2010a). The effect of format on parents' understanding of the risks and benefits of clinical research: A comparison between text, tables, and graphics. *Journal of Health Communication*, *15*, 487–501.

*Tait, A. R., Voepel-Lewis, T., Zikmund-Fisher, B. J., & Fagerlin, A. (2010b). Presenting research risks and benefits to parents: Does format matter? *Anesthesia and Analgesia*, *111*, 718–723.

Trevena, L., Zikmund-Fisher, B., Edwards, A., Timmermans, D., Peters, E., Lipkus, I., . . . Han, P. (2012). Presenting probabilities. In R. Volk & H. Llewellyn-Thomas (Eds.), *Update of the International Patient Decision Aids Standards (IPDAS) Collaboration's background document* (chap. C). Retrieved from http://ipdas.ohri.ca/resources.html

Visschers, V. H., Meertens, R. M., Passchier, W. W., & De Vries, N. N. (2009). Probability information in risk communication: A review of the research literature. *Risk Analysis*, *29*, 267–287.

Vogt, F., & Marteau, T. M. (2012). Perceived effectiveness of stop smoking interventions: Impact of presenting evidence using numbers, visual displays, and different timeframes. *Nicotine & Tobacco Research*, *14*, 200–208.

von Wagner, C., Steptoe, A., Wolf, M. S., & Wardle, J. (2009). Health literacy and health actions: A review and a framework from health psychology. *Health Education & Behavior*, *36*, 860–877.

Waldrop-Valverde, D., Jones, D. L., Gould, F., Kumar, M., & Ownby, R. L. (2010). Neurocognition, health-related reading literacy, and numeracy in medication management for HIV infection. *AIDS Patient Care and STDs*, *24*, 477–484.

Waldrop-Valverde, D., Jones, D. L., Jayaweera, D., Gonzalez, P., Romero, J., & Ownby, R. L. (2009). Gender differences in medication management capacity in HIV infection: The role of health literacy and numeracy. *AIDS and Behavior*, *13*, 46–52.

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.

Weinfurt, K. P., Castel, L. D., Li, Y., Sulmasy, D. P., Balshem, A. M., Benson, A. B., . . . Slater, E. F. (2003). The correlation between patient characteristics and expectations of benefit from phase I clinical trials. *Cancer*, *98*, 166–175.

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making*, *26*, 198–212.

*Woller-Carter, M. M., Okan, Y., Cokely, E. T., & Garcia-Retamero, R. (2012). Communicating and distorting risks with graphs: An eye-tracking study. In *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting* (pp. 1723–1727). Santa Monica, CA: Human Factors and Ergonomics Society.

*Zikmund-Fisher, B. J., Dickson, M., & Witteman, H. O. (2011). Cool but counterproductive: Interactive, web-based risk communications can backfire. *Journal of Medical Internet Research*, *13*, e60.

*Zikmund-Fisher, B. J., Fagerlin, A., & Ubel, P. A. (2010). A demonstration of "less can be more" in risk graphics. *Medical Decision Making*, *30*, 661–671.

Zikmund-Fisher, B. J., Mayman, G., & Fagerlin, A. (2014). Patient numeracy: What do patients need to recognize, think, or do with health numbers? In B. L. Anderson & J. Schulkin (Eds.), *Numerical reasoning in judgments and decision making about health* (pp. 80–104). Cambridge, UK: Cambridge University Press.

Zikmund-Fisher, B. J., Smith, D. M., Ubel, P. A., & Fagerlin, A. (2007). Validation of the subjective numeracy scale: Effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision Making*, *27*, 663–671.

Zikmund-Fisher, B. J., Ubel, P. A., Smith, D. M., Derry, H. A., McClure, J. B., Stark, A., . . . 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.

*Zikmund-Fisher, B. J., Witteman, H. O., Dickson, M., Fuhrel-Forbis, A., Kahn, V. C., Exe, N. L., . . . Fagerlin, A. (2014). Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. *Medical Decision Making*, *34*, 443–453.

*Zikmund-Fisher, B. J., Witteman, H. O., Fuhrel-Forbis, A., Exe, N. L., Kahn, V. C., & Dickson, M. (2012). Animated graphics for comparing two risks: A cautionary tale. *Journal of Medical Internet Research*, *14*, e106.

Rocio Garcia-Retamero is an associate professor of psychology at the University of Granada (Spain) and affiliated scientist at the Max Planck Institute for Human Development (2006–present). She specializes in risk perception and risk communication in health, with emphasis in the design of transparent information formats (e.g., visual aids) to maximize risk understanding across diverse populations (e.g., vulnerable groups and general community samples from several countries). She is also interested in individual differences in abilities (numeracy and graph literacy) and their influence on informed medical decision making. She is cofounder of the multinational RiskLiteracy.org research collaboration.

Edward T. Cokely serves as Presidential Research Professor and associate professor of psychology at the University of Oklahoma and the National Institute for Risk & Resilience. He is an expert in the psychology of skilled and informed decision making, with applications in risk communication and user-friendly technology. He is also affiliated scientist at the Max Planck Institute for Human Development and cofounder of the multinational RiskLiteracy.org research collaboration.

Date received: January 21, 2016
Date accepted: December 23, 2016