Enhancing Understanding and Recall of Quantitative Information about Medical Risks: A Cross-Cultural Comparison between Germany and Spain

Rocio Garcia-Retamero[1], Mirta Galesic[2], and Gerd Gigerenzer[2]

[1]Universidad de Granada (Spain)
[2]Max Planck Institute for Human Development (Germany)

In two experiments, we analyzed cross-cultural differences in understanding and recalling information about medical risks in two countries—Germany and Spain—whose students differ substantially in their quantitative literacy according to the 2003 Programme for International Student Assessment (PISA; OECD, 2003, 2010). We further investigated whether risk understanding can be enhanced by using visual aids (Experiment 1), and whether different ways of describing risks affect recall (Experiment 2). Results showed that Spanish students are more vulnerable to misunderstanding and forgetting the risk information than their German counterparts. Spanish students, however, benefit more than German students from representing the risk information using ecologically rational formats—which exploit the way information is represented in the human mind. We concluded that our results can have important implications for clinical practice.
Keywords: medical risks, Programme for International Student Assessment (PISA), risk perception, risk communication.

En dos experimentos, hemos analizado si existen diferencias culturales en la comprensión y recuerdo de información sobre riesgos médicos. Nos hemos centrado en dos países—Alemania y España—ya que, según los resultados encontrados en el Programme for International Student Assessment (PISA; OECD, 2003, 2010). los estudiantes procedentes de los mismos difieren substancialmente en sus habilidades para procesar información cuantitativa. Así mismo hemos investigado si es posible mejorar la comprensión de los riesgos médicos mediante el uso de material visual de apoyo (Experimento 1), o del uso de diferentes formatos verbales para describir dichos riesgos (Experimento 2). Los resultados de nuestros experimentos han puesto de manifiesto que los estudiantes españoles son más vulnerables a interpretar incorrectamente y a olvidar la información sobre los riesgos que los estudiantes alemanes. Sin embargo, los primeros se benefician en mayor medida que los segundos de la representación de la información sobre riesgos médicos a través del uso de formatos ecológicos—los cuales representan la información de un modo similar a como lo hace la mente humana. Concluimos que nuestros resultados pueden tener implicaciones importantes para la práctica clínica.
Palabras clave: riesgos médicos, Programme for International Student Assessment (PISA), percepción de riesgos, comunicación de riesgos.

Many people have difficulties when processing numerical information and probability expressions (Galesic & Garcia-Retamero, 2010; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, & Woloshin, 2008). Understanding and recalling percentages associated with medical risks and chances of benefits and drawbacks of different treatments, however, is crucial to making informed decisions about health (Feixas & Saul, 2004; Garcia-Retamero & Galesic, 2009b; Gigerenzer & Hoffrage, 1995). Solving problems that require the use of numerical information is broadly called quantitative literacy (Montori & Rothman, 2005). In this paper, we focus on cross-cultural differences in understanding and recalling medical information between two countries whose citizens differ substantially in their quantitative literacy: Germany and Spain.

The Programme for International Student Assessment (PISA) is a system of international assessments carried out every 3 years by the Organization for Economic Cooperation and Development (OECD) to measure 15-year-old students' literacy skills. Quantitative literacy was the subject area assessed in the 2003 study, which focused on four broad areas: having a basic number sense, recognizing geometric patterns, understanding functional relationships, and understanding the concept of uncertainty (OECD, 2003; see also OECD, 2010).

The 2003 PISA study results show that overall performance of Spanish students in quantitative literacy was substantially lower than the average performance for students from most OECD countries (484 vs. 500 points). Spanish students also performed below the OECD average on each quantitative literacy subscale representing a specific content area. The performance of German students in the combined quantitative literacy score (i.e., 503 points) and in all subscales was substantially higher than that of Spanish students and also higher than the average performance of students from most OECD countries. Along with scale scores, the 2003 PISA study also used six proficiency levels (with level 6 being the highest) to describe student performance in quantitative literacy. Spain had a greater percentage of students at level 1 than Germany (15 vs. 12%), and a lower percentage at level 6 (1 vs. 4%). Results for each of the four quantitative content areas followed a similar pattern. In Spain, with an average score of 597, even the highest achievers (those in the top 10%) were outperformed on average by their OECD counterparts (including Germans, with an average score of 632). The lower performance of students in Spain than in Germany was shown not only in the PISA 2003 study, but also in other studies that measure mathematical skills, such as the Trends in Mathematics and Science Study (TIMSS), whose results are highly correlated with those of the 2003 PISA study ($r = .91$; Rindermann, 2007).

Cultural and material resources are important predictors of students' achievement in PISA studies (Marks, Cresswell, & Ainley, 2006). There is a strong relationship between national wealth (i.e., gross domestic product per capita) and the results of PISA studies (Hunt & Wittmann, 2008; Rindermann, 2007). That is, the economic status of a country explains a significant amount of the variance in the measure of cognitive competence of that country's residents. It is likely that wealthier nations, on a per capita basis, can provide better schools, better health care, and more stable living conditions for students, and that these conditions improve mathematical competence (Hunt & Wittmann, 2008; Marks et al., 2006). In fact, there are important differences in education policies between Germany and Spain, which could explain the emerging cross-cultural differences in mathematical performance of their students (Eurostat, 2007, 2009). For instance, in 2009, the public and private overall expenditure per capita on education in Germany was larger than that in Spain. The annual expenditure on educational institutions per student was also noticeably higher. In 2009, Germany had a higher percentage of young people attaining secondary- and tertiary-level education (especially in science, math, and computing), and a much higher student participation in formal and informal learning activities. Germany also had a lower rate of early school leavers and people with no formal education (Instituto Nacional de Estadística, 2009; Statistisches Bundesamt Deutschland, 2009).

The first question we address in this paper is whether there are cross-cultural differences in understanding and recalling information about medical risks between students in Germany and Spain. There is a dearth of published research on differences between countries in risk perception and medical decision making (Garcia-Retamero & Galesic, in press). To the best of our knowledge, cross-cultural comparisons between Germany and Spain on this issue are lacking.

The second question we address is whether people—especially those who lag behind in general understanding of numerically presented data—can be helped to understand and recall quantitative information about health. A wide range of aids have been proposed as potentially promising methods for communicating medical risks (Garcia-Retamero & Galesic, 2009a; Garcia-Retamero, Galesic, & Gigerenzer, 2010). Some of these aids are considered *ecologically rational* because they are formatted to exploit the way information is represented in the human mind (Gigerenzer et al., 2008; Gigerenzer, Todd, & the ABC Research Group, 1999). Visual aids and descriptions of consequences of risks linked to people's everyday experiences are two prominent examples of ecologically rational formats.

Visual aids such as icon arrays[1] can improve understanding of risks and benefits associated with different treatments,

---

[1] Graphical representations consisting of a number of circles or other icons symbolizing individuals who are affected by some risk.

screenings, and life-styles (Ancker, Senathirajah, Kukafka, & Starren, 2006; Galesic & Garcia-Retamero, 2010b, in press; Lipkus, 2007; Lipkus & Hollands, 1999; Paling, 2003). They can also promote consideration of beneficial treatments that have side effects (Waters, Weinstein, Colditz, & Emmons, 2007), limit errors induced by anecdotal narratives (Fagerlin, Wang, & Ubel, 2005), and framed messages (Garcia-Retamero & Galesic, 2010a), and reduce biases (Garcia-Retamero & Dhami, 2011; Garcia-Retamero & Galesic, 2009a; Garcia-Retamero et al., 2010). Research by Galesic, Garcia-Retamero, and Gigerenzer (2009; see also Garcia-Retamero & Galesic, 2010b) showed that visual aids are useful when the risk information is presented both in complex numerical formats such as relative risk reduction (e.g., aspirin can reduce the risk of having a stroke or heart attack by 13% in people with symptoms of arterial disease) and in numerical formats that are relatively easy to understand (absolute risk reduction; e.g., aspirin can reduce the risk of having a stroke or heart attack from 8% to 7% in people with symptoms of arterial disease). Visual aids therefore lead to additional improvements in accuracy even when transparent numerical representations are used. More importantly, these aids can be very helpful for older people and those with limited numeracy and language proficiency (Galesic et al., 2009; Garcia-Retamero & Galesic, 2009a; Garcia-Retamero et al., 2010).

Similarly, previous research by Galesic and Garcia-Retamero (2011) showed that the way that consequences of risky behaviors are described influences people's recall substantially. Doctors and health authorities, for instance, usually communicate consequences of risky life-styles in terms of increased risk of different diseases. As Kenchaiah et al. (2002; see also Miller, Balady, & Fletcher, 1997) pointed out if a person does not exercise regularly, her risk of developing a cardiovascular disease is expected to be increased by one third. This information is often difficult to understand and recall–especially for people with limited numerical skills because they have problems with mapping the concept "risk of disease" to concrete instances from everyday experiences (Galesic & Garcia-Retamero, 2011). Consequences of risky behaviors can also be expressed in less abstract terms, such as changes in "life years," a concept that is more natural to most people, as everybody is exposed to the experience of getting older and watching others grow old and die (Kenchaiah et al., 2002; Miller et al., 1997). As Galesic and Garcia-Retamero (2011) showed, when information about consequences of risky behaviors is presented as years of life lost or gained, recall is much better than when it is presented in terms of risk of a disease. This is so for participants with high and low numerical skills.

Yet our understanding of the effectiveness of these ecologically rational formats in enhancing understanding and recall of information about medical risks in different cultures remains incomplete. In two experiments, we investigated the effectiveness of these formats in a between-countries comparison. The experiments were conducted in

our laboratories at the Max Planck Institute for Human Development in Berlin (Germany) and at the Department of Experimental Psychology of the University of Granada (Spain), respectively. In Experiment 1, we studied how helpful visual aids are for understanding treatment risk reduction when the risk information is presented either as absolute or relative risk reduction. In Experiment 2, we investigated whether consequences of risky behaviors are better recalled when they are expressed as changes in life expectancy rather than as changes in risk of diseases.

Experiment 1: Do Visual Aids Enhance Understanding of Risk Reduction?

In a cross-cultural study, we tested whether German and Spanish students differ in their ability to understand medical information. First following Schwartz, Woloshin, Black, and Welch (1997), we measured accuracy of students' estimates of treatment risk reduction. Second, we analyzed whether visual aids are especially helpful for communicating risk reduction in the population with lower quantitative literacy. Finally, we investigated whether visual aids lead to additional improvements in accuracy even if transparent numerical formats such as absolute risk reduction are used.

Method

Participants

The participants were 230 undergraduate students: 118 from the University of Granada (Spain; 48 men and 70 women) and 112 from the Free University of Berlin (Germany; 49 men and 63 women). The two universities are comparable in terms of size, courses of study offered, and type of students they attract. Participants specified their nationality: All participants in Spain and Germany were Spanish and German, respectively. German and Spanish participants had a median age of 20 (range 18–42) and 24 years (range 18–34), respectively. Participants were invited to the laboratory and received 10 euros. They were randomly assigned to the experimental groups.

Procedure

All participants completed a 25-min computerized questionnaire. The questionnaire was developed in English and translated into German and Spanish by skilled translators. The materials were back-translated to ensure comparability and, therefore, were comparable. All translations were performed by skilled translators. When programming the questionnaire, special care was taken to ensure the interface looked the same in the German and Spanish versions. The Ethics Committee of the

Max Planck Institute for Human Development approved the methodology of the study. At the beginning of the experiment, all participants gave their consent.

The participants were presented with three medical scenarios: one involving the usefulness of aspirin in reducing the risk of stroke or heart attack for people with symptoms of arterial disease; another about a hypothetical new drug (modeled after the drug Simvastatin; Skolbekken, 1998) for reducing cholesterol that also decreases the risk of dying from a heart attack; and the third about a fictitious case of screenings for early signs of acute appendicitis. The order of the scenarios was randomized.

Two independent variables were manipulated between subjects. First, half of the participants got the numerical information in the form of either absolute or relative risk reduction. For instance, in the first scenario, participants got the following text in the relative risk reduction condition: "For people with symptoms of arterial disease, aspirin can reduce the risk of having a stroke or heart attack by 13%." Those in the absolute risk reduction condition received "For people with symptoms of arterial disease, aspirin can

reduce the risk of having a stroke or heart attack: 8% of such people who did not take aspirin had a stroke or heart attack, compared to 7% of such people who did take aspirin." In the second scenario, participants were told that the drug can reduce the risk of dying from a heart attack by 38% (relative risk reduction) or from 8% to 5% (absolute risk reduction). Finally, in the third scenario, participants were told that the screening can reduce the risk of acute appendicitis by 60% (relative risk reduction) or from 8% to 3% (absolute risk reduction). For simplicity in what follows, we refer to the consequences—stroke, heart attack, appendicitis—as "illness."

Independently of this manipulation, half of the participants got—in addition to the numerical information—two icon arrays, one presenting baseline risk of the disease without treatment (or screening), and the other presenting the risk with treatment (or screening). All icon arrays contained 1,000 circles. Affected individuals were shown as black circles at the end of the array. We used circles because previous research (Stone, Yates, & Parker, 1997) did not find differences in effects of arrays with faces compared
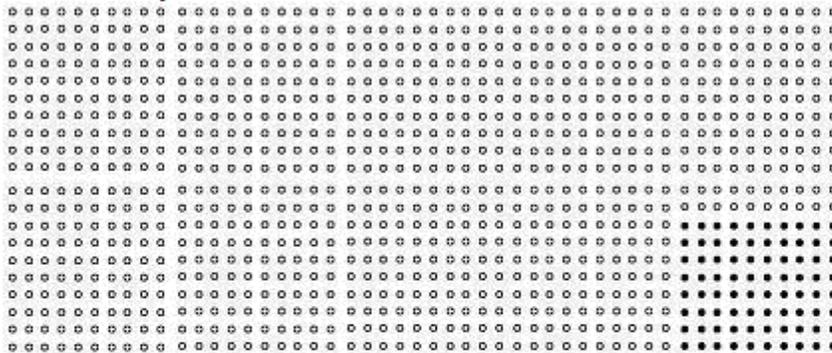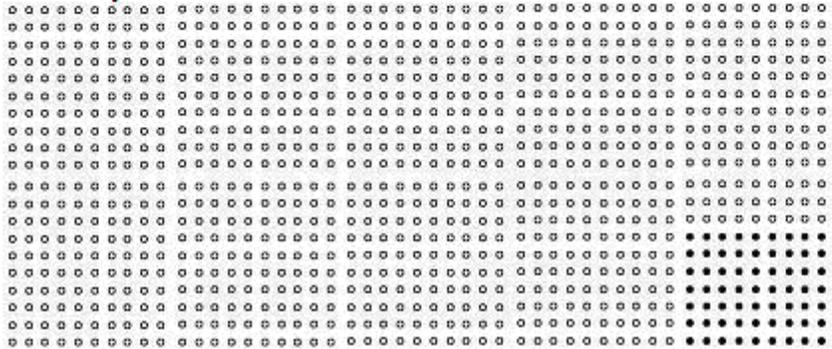

Figure 1. Example of a condition from Experiment 1: Numerical information about relative risk reduction involving icon arrays.

to more abstract symbols. An example of the condition involving icon arrays is shown in Figure 1 (original material was in either German or Spanish).

After reading the information about each medical scenario, participants were asked how many of 1,000 people might become ill if they *do not* take the drug (or participate in screening). Afterward, they were asked how many of 1,000 people might become ill if they *do* take the drug (or go to screening). In addition, participants estimated, on scales from 1 (not helpful at all) to 15 (very helpful), how helpful the treatment was for preventing illness, and how helpful the numerical information about the treatment was for answering the questions.

Statistical analyses. Following Schwartz et al. (1997), participants' accuracy of risk understanding was assessed from their responses to the first two questions. By deducting the second answer from the first and dividing it by the first answer, we calculated the estimated relative risk reduction and coded it as correct if it was within ±3 percentage points of the right answer. This criterion was determined based on the distribution of errors, which peaked within this interval and quickly became flat after passing its limits. Following Galesic et al. (2009; see also Garcia-Retamero & Galesic, 2010b), we computed the number of participants who gave an accurate estimate of relative risk reduction in at least one of the scenarios.

As there were no differences in effects of presentation format between the different scenarios (aspirin, cholesterol drug, and appendicitis screening), we analyzed them together. To test the effects of participants' nationality (Spanish vs. German), information format (relative vs. absolute risk reduction), and icon arrays (presence vs. absence), we conducted analyses of variance (ANOVAs), with the number of participants who estimated relative risk reduction correctly in at least one scenario, and participants' estimation of helpfulness of the treatment and helpfulness of the information as dependent variables. We followed Lunney (1970; see also Cleary & Angel, 1984), who showed that ANOVAs can be used to obtain conservative results for large samples of a dichotomous dependent variable.

Results

Understanding numerical information about risk reduction was more difficult for Spanish than for Germans students. The percentage of German students who gave an accurate estimate of relative risk reduction in at least one of the scenarios was higher than that of Spanish students, 87% versus 53%; $F(1,222) = 40.37$, $p = .001$, $\eta_p^2 = .18$. Both numerical format and icon arrays had a positive effect on accuracy in German and Spanish students. Specifically, large improvements in accuracy were achieved when numerical information was presented in terms of absolute rather than relative risk reduction, 83% versus 56%, $F(1,222) = 25.47$, $p = .001$, $\eta_p^2 = .11$. Icon arrays were a useful addition to both types of numerical representation, 80% versus 60%, $F(1,222) = 13.7$, $p = .002$, $\eta_p^2 = .10$. More interestingly, icon arrays were particularly useful for lower skilled participants
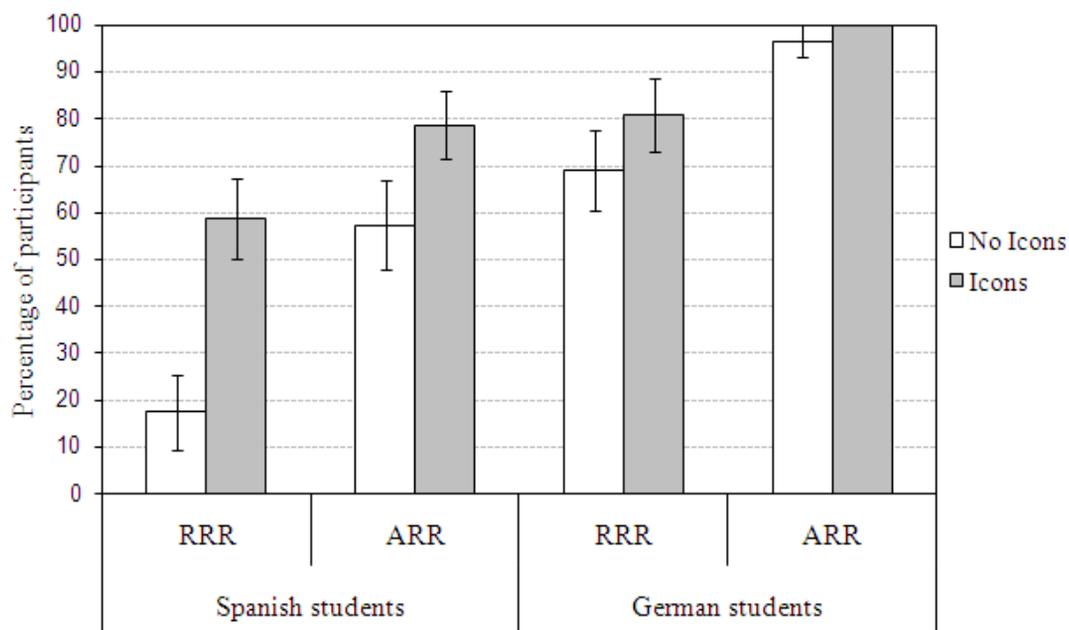

Figure 2. Percentage of Spanish and German students who gave a correct estimate of relative risk reduction in at least one of the scenarios when presented as absolute (ARR) or relative (RRR) risk reduction involving icon arrays.

both when the risk information was presented in the form of absolute and relative risk reduction, $F(1,222) = 5.14$, $p = .024$, $\eta^2_p = .10$ (see Figure 2).

Both Spanish and German students perceived that the treatment was more helpful for preventing illness when the numerical information was presented in terms of relative instead of absolute risk reduction, 9.6 versus 7.8; $F(1,222) = 17.34$, $p = .001$, $\eta^2_p = .20$. In addition, the information provided for answering the questions was perceived to be more useful when icon arrays were given in addition to the numerical information than when only numerical information was presented, 9.7 versus 8.9; $F(1,222) = 3.98$, $p = .007$, $\eta^2_p = .10$.

Conclusions

In this study, we analyzed whether there are cross-cultural differences in understanding numerical information about risk reduction between students in Germany and Spain, and whether icon arrays enhance understanding when the risk information is presented in the form of either absolute or relative risk reduction. Results show that Spanish students had more problems understanding the medical risks than their German counterparts. Using icon arrays, however, was a useful method for communicating treatment risk reduction both when the risk information was presented in absolute and in relative terms.

Experiment 2: Do Different Ways of Describing Consequences of Risky Behaviors Affect Recall?

The second cross-cultural study was designed to test whether Spanish students would also have more difficulties recalling long-term consequences of risky behaviors than their German counterparts. We further analyzed how different ways of describing these consequences (i.e., increase or decrease of life expectancy vs. risk of disease) affect recall in students from both countries.

Method

Participants

Two hundred and thirty undergraduate students participated in the experiment (118 from the University of Granada, Spain; 48 men and 70 women; and 112 from the Free University of Berlin, Germany; 49 men and 63 women. Participants specified their nationality: All participants in Spain and Germany were Spanish and German, respectively. German and Spanish participants had a median age of 20 (range 18–42) and 24 years (range 18–34), respectively. Participants were invited to the laboratory twice and received 10 euros per session. They were randomly assigned to the experimental groups.

Procedure

The study was conducted in two waves 2 weeks apart. In the first wave, all participants completed a computerized questionnaire. In the second wave, they were asked to recall the information that they had read in the first wave. As in Experiment 1, the questionnaire was developed in English and translated into German and Spanish. When programming the questionnaire, special care was taken to ensure the interface looked the same in the German and Spanish versions. The Ethics Committee of the Max Planck Institute for Human Development approved the methodology of the study. At the beginning of the experiment, all participants gave their consent.

In the first wave of the study, the participants were presented with two medical scenarios involving realistic risks related to obesity (Fontaine, Redden, Wang, Westfall, & Allison, 2003; Franco et al., 2005) and physical inactivity (Kenchaiah et al., 2002; Miller et al., 1997), respectively. One group of participants got the information in terms of an increase or decrease in risk of heart failure. Specifically, these participants were told: "If you are overweight, you are expected to have two times greater risk of heart failure than if your weight is normal" and "if you exercise regularly, your risk of developing a cardiovascular disease is expected to be reduced by about one third." The other group got the information in terms of an increase or decrease in life expectancy: "If you are overweight, you are expected to live about 5 years less than if your weight is normal" and "if you exercise regularly, you are expected to live about 3 years more than if you do not exercise." Participants were not instructed to memorize this information.

Overweight participants ($n = 62$; as determined by questions about their height and weight) estimated on scales from 1 (very unlikely) to 15 (very likely) how likely it was that they would try to reduce their weight and exercise more often after reading the information. In addition, after completing unrelated questions about health risks, which took approximately 45 min, all participants were asked to recall the information presented in the two scenarios. Specifically, participants who got the information in terms of expected risk of heart failure answered the question: "How much higher is the risk of heart failure for people who are overweight, compared to an average person?" and the question: "How much lower is the risk of cardiovascular disease for people who exercise regularly, compared to an average person?" Participants who got the information in terms of life expectancy answered the question: "How much shorter is life expectancy for people who are overweight, compared to an average person?" and the question: "How much longer is life expectancy for people who exercise

regularly, compared to an average person?" The same procedure was repeated in the second step, conducted after 2 weeks. The answers were coded as accurate if they were exactly correct.

Statistical analyses

Following Galesic and Garcia-Retamero (2011), we conducted ANOVAs with the number of participants who recalled the medical information correctly in at least one scenario as a dependent variable, the nationality (Spanish vs. German) and the information format (changes in risk of disease vs. changes in life expectancy) as between-subjects factors, and the recall delay (after 45 min vs. after 2 weeks) as a within-subject factor. We also conducted ANOVAs with participants' intentions to change behaviors after reading the information as a dependent variable, the nationality and the information format as between-subjects factors, and the task (obesity and physical inactivity) as a within-subjects factor on overweight participants.

Results

When numerical information was presented as years of life lost or gained recall was better than when it was presented in terms of more abstract risks, 50% versus 78%, $F(1,222) = 39.47$, $p = .001$, $\eta^2_p = .26$. The improvements

in recall when presented in terms of number of years were particularly large for Spanish students both after 45 min and after 2 weeks (see Figure 3), $F(1,222) = 9.05$, $p = .002$, $\eta^2_p = .10$. Furthermore, both Spanish and German overweight students were more likely to indicate they would try to lose weight or change their current exercise behavior when the information was presented as years of life lost or gained than when it was presented as an abstract increase in risk, 9.17 versus 11.82, $F(1,58) = 11.24$, $p = .001$, $\eta^2_p = .20$.

Comment

In two experiments, we analyzed cross-cultural differences in understanding and recalling information about medical risks between two countries—Germany and Spain—whose students differ substantially in their quantitative literacy (OECD, 2003, 2010; Rindermann, 2007). We also studied whether understanding and recalling can be enhanced in these populations.

Results show that Spanish students were more vulnerable to misunderstanding and forgetting the risk information than their German counterparts. This finding suggests that the former students were less skilled at processing not only numerically presented data but also health statistics. Our results further show that especially Spanish students can be aided by using ecologically rational formats to enhance comprehension and recall. The largest
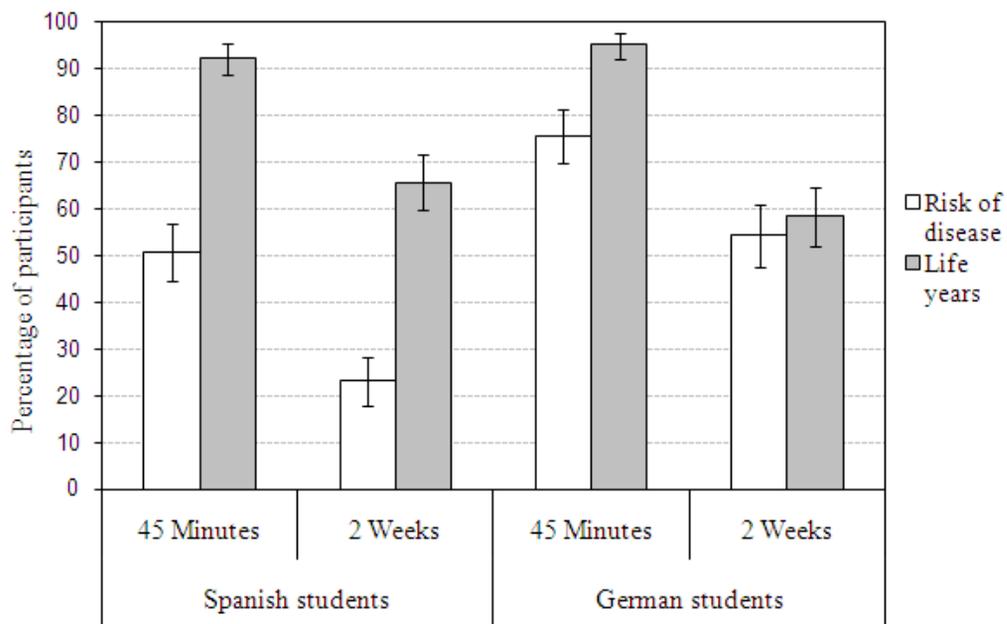

Figure 3. Percentage of Spanish and German students who correctly recalled the medical information in at least one scenario when presented in terms of risk of heart disease or in terms of life years gained or lost after 45 min or 2 weeks.

improvement in risk understanding was shown when Spanish participants received the numerical information as relative risk reduction–a representation format that is less transparent than absolute risk reduction (Covey, 2007; Garcia-Retamero & Galesic, 2010)–with icon arrays accompanying the numerical descriptions.

In a similar vein, describing consequences of risky behaviors as reductions in life expectancy rather than as increments in risk of heart disease improved recall. Again, the improvement was particularly large for Spanish students, both short and long term. In fact, presenting consequences of risky behaviors as a reduction in "life years" raised Spanish students' understanding to the level of German students'.

Cross-cultural differences between students in Germany and Spain are important in light of the available information about health and medical risks in these countries. For instance, according to the World Health Organization (2009), specialized health-related magazines and brochures are relatively rare in Spain compared to Germany. Similarly, while independent health-related websites are relatively scarce in Spain, Germany has a wide range of web-based health support systems, often sponsored by the government or by insurance providers (Eurostat, 2009). Spaniards might then be less proactive in seeking information about health on their own compared to the citizens of Germany. Additionally, due to their limitations in understanding and recalling information about health, Spaniards might also have more difficulties when making their own medical decisions. Ecologically rational formats might then be especially suitable for this population and their use should be strongly recommended.

Our results are consistent with previous research showing that problems with understanding and recalling numerical information often do not reside in people's minds–but in the representation of the problem (Elmore & Gigerenzer, 2005; Gigerenzer & Edwards, 2003; Gigerenzer et al., 2008; Gigerenzer & Hoffrage, 1995). Our findings also have important implications for medical practice as they suggest suitable ways to communicate quantitative medical data to people who are vulnerable to having difficulty when making decisions about health. These results are also promising for public health campaigns aimed at improving people's awareness about unhealthy consequences of risky life-styles, such smoking, obesity, and physical inactivity.

A limitation of our experiments is that they were conducted in a laboratory setting and did not involve real patient–doctor interactions (Garcia-Retamero & Galesic, 2009c). Although this controlled setting enabled us to draw clearer conclusions about the effects of various properties of ecologically rational formats, it is possible that the formats tested would show additional benefits in clinical settings (e.g., increase patients' trust in physicians and willingness to share decision making; Garcia-Retamero & Galesic, 2009c; Garcia-Retamero, Takezawa, & Gigerenzer, 2009). Further research could investigate this issue.

References

Ancker, J. S., Senathirajah, Y., Kukafka, R., & Starren, J. B. (2006) Design features of graphs in health risk communication: A systematic review. *Journal of American Medical Informatics Association*, *13*, 608-618. doi:10.1197/jamia.M2115

Cleary, P. D., & Angel, R. (1984). The analysis of relationships involving dichotomous dependent-variables. *Journal of Health and Social Behavior*, *25*, 334-348. doi:10.2307/2136429

Covey, J. (2007). A meta-analysis of the effects of presenting treatment benefits in different formats. *Medical Decision Making*, *27*, 638-654. doi:10.1177/0272989X07306783

Elmore, J. G., & Gigerenzer, G. (2005). Benign breast disease-the risks of communicating risk. *New England Journal of Medicine*, *353*, 297-299. doi:10.1056/NEJMe058111

Eurostat. (2007). *Europe in figures: Education*. Retrieved from: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-CD-06-001-02/EN/KS-CD-06-001-02-EN.PDF.

Eurostat. (2009). *Europe in figures: Eurostat year book 2009*. Retrieved from: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-CD-09-001/EN/KS-CD-09-001-EN.PDF.

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. doi:10.1177/0272989X05278931

Feixas, G., & Saul, L. A. (2004). The multi-center dilemma project: An investigation of the role of cognitive conflicts in health. *The Spanish Journal of Psychology*, *7*, 69-78.

Fontaine, K. R., Redden, D. T., Wang, C., Westfall, A. O., & Allison, D. B. (2003). Years of life lost due to obesity. *The Journal of the American Medical Association*, *289*, 187-193. doi:10.1001/jama.289.2.187

Franco, O. H., Laet, C. D., Peeters, A., Jonker, J., Mackenbach, J., & Nusselder, W. (2005). Effects of physical activity on life expectancy with cardiovascular disease. *Archives of Internal Medicine*, *165*, 2355-2360. doi:10.1001/archinte.165.20.2355

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. doi:10.1001/archinternmed.2009.481

Galesic, M., & Garcia-Retamero, R. (2011). Communicating consequences of risky behaviors: Life expectancy versus risk of disease. *Patient Education and Counseling, 82,* 30-35. doi:10.1016/j.pec.2010.02.008

Galesic, M., & Garcia-Retamero, R. (in press). Graph literacy: A cross-cultural comparison. *Medical Decision Making*.

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

Garcia-Retamero, R., & Galesic, M. (2009a).Communicating treatment risk reduction to people with low numeracy skills:

A cross-cultural comparison. *American Journal of Public Health*, *99*, 2196-2202. doi:10.2105/AJPH.2009.160234

Garcia-Retamero, R., & Galesic, M. (2009b). Heuristics. In M. W. Kattan (Ed.), *The encyclopedia of medical decision making* (pp. 596-599). Thousand Oaks, CA: SAGE Publications.

Garcia-Retamero, R., & Galesic, M. (2009c). Trust in health care. In M. W. Kattan (Ed.), *The encyclopedia of medical decision making* (pp. 1153-1155). Thousand Oaks, CA: SAGE Publications.

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. doi:10.1007/s11606-010-1484-9

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. doi:10.1016/j.socscimed.2010.03.019

Garcia-Retamero, R., & Galesic, M. (in press). *Risk communication and medical decision making: A cross-cultural perspective*. 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. doi:10.1177/0272989X10369000

Garcia-Retamero, R., Takezawa, M., & Gigerenzer, G. (2009). Does imitation benefit cue order learning? *Experimental Psychology*, *56*, 307-320. doi:10.1027/1618-3169.56.5.307

Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal*, *327*, 741-744. doi:10.1136/bmj.327.7417.741

Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., & Woloshin, S. (2008). Helping doctors and patients make sense of health statistics**. ***Psychological Science in the Public Interest*, *8***, 53-96.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684-704. doi:10.1037//0033-295X.102.4.684

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.

Hunt, E., & Wittmann, W. (2008). National intelligence and national prosperity. *Intelligence*, *36*, 1-9. doi:10.1016/j.intell.2006.11.002

Instituto Nacional de Estadística [National Institute of Statistics]. (2009). *Cifras de población y censo demográfico* [population figures and demographic census]. Retrieved from: http://www.ine.es/inebmenu/mnu_cifraspob.htm.

Kenchaiah, S., Evans, J. C., Levy, D., Wilson, P. W. F., Benjamin, E. J., Larson, M. G.,… Vasan, R. S. (2002). Obesity and the risk of heart failure. *New England Journal of Medicine*, *347*, 305-313. doi:10.1056/NEJMoa020245

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risk: Suggested best practices and future recommendations. *Medical Decision Making*, *27*, 696-713. doi:10.1177/0272989X07307271

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

Lunney, G. H. (1970). Using analysis of variance with a dichotomous dependent variable–An empirical study. *Journal of Educational Measurement*, *7*, 263-269. doi:10.1111/j.1745-3984.1970.tb00727.x

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation*, *12*, 105-128. doi:10.1080/13803610600587040

Miller, T. D., Balady, G. J., & Fletcher, G. F. (1997). Exercise and its role in the prevention and rehabilitation of cardiovascular disease. *Annals of Behavioral Medicine*, *3*, 220-229. doi:10.1007/BF02892287

Montori, V. M., & Rothman, R. L. (2005). Weakness in numbers–The challenge of numeracy in healthcare. *Journal of General Internal Medicine*, *11*, 1071-1072. doi:10.1111/j.1525-1497.2005.051498.x

OEDC (2003). *The PISA 2003 Assessment framework–Mathematics, reading, science, and problem solving knowledge and skills*. Paris: OEDC.

OECD (2010). *PISA 2009 Results: What students know and can do – Student performance in reading, mathematics and science (Volume I)*. Paris: OEDC. doi.org/10.1787/9789264091450-en

Paling, J. (2003). Strategies to help patients understand risks. *British Medical Journal*, *327*, 745-748. doi:10.1136/bmj.327.7417.745

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, *21,* 667-706. doi:10.1002/per.634

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.

Skolbekken, J. A. (1998). Communicating the risk reduction achieved by cholesterol reducing drugs. *British Medical Journal*, *316*, 1956-1958.

Statistisches Bundesamt Deutschland. (2009). *Microcensus*. Retrieved from: https://www-genesis.destatis.de/genesis/online.

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. doi:10.1037//1076-898X.3.4.243

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. doi:10.1037/1076-898X.13.1.11

World Health Organization. (2009). *Core health indicators*. Retrieved from: http://www.who.int/whosis/database/core/core_select_process.cfm?countries=all&indicators=nha.