Blocks, Ovals, or People? Icon Type Affects Risk Perceptions and Recall of Pictographs

Brian J. Zikmund-Fisher, PhD, Holly O. Witteman, PhD, Mark Dickson, MA, Andrea Fuhrel-Forbis, MA, Valerie C. Kahn, MPH, Nicole L. Exe, MPH, Melissa Valerio, PhD, Lisa G. Holtzman, MPH, Laura D. Scherer, PhD, Angela Fagerlin, PhD

Background. *Research has demonstrated that icon arrays (also called ''pictographs'') are an effective method of communicating risk statistics and appear particularly useful to less numerate and less graphically literate people. Yet research is very limited regarding whether icon type affects how people interpret and remember these graphs.* ***Methods.*** *1502 people age 35–75 from a demographically diverse online panel completed a cardiovascular risk calculator based on Framingham data using their actual age, weight, and other health data. Participants received their risk estimate in an icon array graphic that used 1 of 6 types of icons: rectangular blocks, filled ovals, smile/frown faces, an outline of a person's head and shoulders, male/female ''restroom'' person icons (gender matched), or actual head-and-shoulder photographs of people of varied races (gender matched). In each icon array, blue icons represented cardiovascular events and gray icons represented those who would not experience an event. We measured perceived risk magnitude, approximate recall, and opinions about the icon arrays, as well as subjective numeracy and an abbreviated measure of graphical literacy.* ***Results.*** *Risk recall was significantly higher with more anthropomorphic icons (restroom icons, head outlines, and photos) than with other icon types, and participants rated restroom icons as most preferred. However, while restroom icons resulted in the highest correlations between perceived and actual risk among more numerate/graphically literate participants, they performed no better than other icon types among less numerate/graphically literate participants.* ***Conclusions.*** *Icon type influences both risk perceptions and risk recall, with restroom icons in particular resulting in improved outcomes. However, optimal icon types may depend on numeracy and/or graphical literacy skills.* ***Key words:*** *risk; patient education as topic; patient-provider communication; decision aids; visual aids* ***(Med Decis Making 2014;34:443–453)***

Icon array risk graphics (sometimes called ''pictographs'') are matrices of icons that represent a potentially at-risk population. Icon arrays are most frequently presented using denominators of 100, but sometimes 1,000 or even 10,000. Within the array, some icons are distinctly colored or shaded to show the number of individuals who experience a particular condition or event (e.g., a diagnosis of cancer or the experience of a side effect), while the remaining figures (often gray) represent those who are not affected. This structure means that a viewer can learn about his or her risk either by focusing on the ''part-whole'' relationship (i.e., the relative area occupied by event v. nonevent icons) or by specifically counting the number of risk events. It also means that both the risk ''loss'' frame (i.e., likelihood of negative events) and ''gain'' frame (i.e., likelihood of positive events) are visually salient.

Over the past 10 years, a growing body of research has provided evidence that icon arrays are a particularly valuable method of communicating risk statistics. Multiple researchers have shown that use of icon arrays supports the acquisition of both gist and precise risk understanding,[1,2] and this is especially true for people with lower numeracy skills.[3,4] Icon arrays also help mitigate biases related to testimonials[5] or varying denominators[6,7] and can reduce side effect aversion.[8]

Yet, despite the attention paid to testing icon arrays versus other graphics, very little attention has been paid to the choice of what kind of icon to use in an array. More than 15 years ago, a study

DOI: 10.1177/0272989X13511706

compared risk perceptions generated from viewing displays using different types of icons and found that icon type had no significant effects.[9] However, this study used very crude displays (e.g., asterisks v. crude stick figures) and only displayed icons for events rather than a full array that would represent both affected and unaffected individuals. To our knowledge, no systematic investigation of the impact of icon type within icon array graphics has ever been conducted.

In the absence of such data, researchers and decision aid developers have used a variety of icon types to display risk statistics. Published studies have used rectangular blocks,[1,2,7,10–13] ovals,[14,15] circles,[3,4,6,16–20] diamond shapes,[21] cartoon smile faces,[22–25] stick figures,[8] head-and-shoulders outlines of people,[5,26] and person icons of the type often used to identify male and female restrooms.[27–32] Whether the choice among these widely varying icon types might have influenced risk perceptions or other study results has remained uninvestigated.

Received 18 June 2013 from Department of Health Behavior and Health Education, University of Michigan, Ann Arbor, MI (BJZF); Division of General Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, MI (BJZF, AF); Center for Bioethics and Social Sciences in Medicine, University of Michigan, Ann Arbor, MI (BJZF, HOW, MD, AFF, VCK, NLE, MV, LGH, LDS, AF); Risk Science Center, University of Michigan, Ann Arbor, MI (BJZF); Office of Education and Continuing Professional Development, Faculty of Medicine, Université Laval, Quebec City, Quebec, Canada (HOW); Department of Family and Emergency Medicine, Faculty of Medicine, Université Laval, Quebec City, Quebec, Canada (HOW); Research Centre of the CHU de Québec, Quebec City, Quebec, Canada (HOW); Sitesteaders Development, Ann Arbor, MI (MD); Research Center for Group Dynamics, Institute for Social Research, University of Michigan, Ann Arbor, MI (AFF); Division of Health Promotion and Behavioral Science, University of Texas School of Public Health at Houston, San Antonio Regional Campus, San Antonio, TX (MV); ReThink Health, Morristown, NJ (LGH); Department of Psychological Sciences, University of Missouri, Columbia, MO (LDS); Department of Psychology, University of Michigan, Ann Arbor, MI (LDS, AF); and Center for Clinical Management Research, VA Ann Arbor Healthcare System, Ann Arbor, MI (AF). This research was supported by a grant from the National Cancer Institute (R01 CA087595, P.A. Ubel, PI). Dr. Zikmund-Fisher was supported by a career development award from the American Cancer Society (MRSG-06-130-01-CPPB). The funding agreements ensured the authors' independence in designing the study, interpreting the data, and publishing the report. The authors have no conflicts of interest to report. Revision accepted for publication 12 October 2013.

Address correspondence to Brian J. Zikmund-Fisher, Department of Health Behavior and Health Education, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109-2029; telephone: (734) 936-9179; fax: (734) 936-8944; e-mail: bzikmund@umich.edu.

To begin to fill this knowledge gap and provide concrete guidance for risk communicators, we conducted a randomized experiment in which participants completed a cardiovascular risk calculator and then received their risk estimates in 1 of 6 icon array formats. We assessed whether varying the icon used in the icon arrays would alter people's risk perceptions, their recall of risk information, or their preferences regarding these graphics. In addition, we assessed participants' numeracy and graphical literacy and explored whether any differences associated with icon type were consistent or different among people with higher versus lower numeracy or graphical literacy skills.

METHODS

Recruitment

A stratified random sample of US adults age 35–75 was selected from a panel of Internet users administered by Survey Sampling International (SSI) (Shelton, Connecticut), which recruits panel members through a variety of opt-in methods. To ensure demographic diversity (although not necessarily representativeness) and offset large expected variations in response rates, we drew distinct subsamples by both age and race (thereby roughly approximating the distributions of these characteristics in the US population), and the number of e-mail invitations in each demographic subsample was dynamically adjusted until all quotas were achieved. Selected panel members received an e-mail invitation with a personalized link to complete the online survey with 1 reminder e-mail for nonresponders. SSI tracked participation via unique identification numbers to prevent duplicate uses of the same link to participate. We recruited for a 2-week period in October 2011. Upon completion, participants were entered into both an instant-win contest and a monthly drawing administered by SSI for modest prizes.

Design of the Study

Study participants completed a cardiovascular risk calculator that implemented the D'Agostino (2008) model developed from Framingham study data.[33] This is a robust model that estimates an individual's risk of cardiovascular disease (CVD) within the next 10 years based on her or his age, blood pressure, cholesterol level or body mass index (BMI), and whether she or he smokes, is diabetic, and takes blood pressure medication. The model generates 10-year risk estimates that range from 1% to ''more than 30%,'' and all estimates were rounded to the nearest whole percentage based on recent research showing that integer risk statistics are perceived as more trustworthy than statistics with decimals.[34] Because the model applies only to people who have not already experienced CVD,[33] we excluded any respondents who reported prior heart disease or stroke. For participants who did not know their blood pressure, we made conservative estimates based on mean blood pressure for their age and whether they stated they took blood pressure medication. All remaining participants were then shown their personal risk both in a large size icon array format and numerically. Throughout the study, rather than using the term ''general cardiovascular disease,'' we described CVD to participants with more familiar terms, ''heart disease and stroke.''

We randomly assigned participants to view icon array displays using 1 of 6 types of icons (Figure 1). We note here that for simplicity we use the term *icon* according to the conventions of health communication research to refer to any image in an icon array. However, in the original language of visual displays of information, the term *icon* refers specifically to a representation of the likeness of an object, an *index* refers to an image that is associated with the object, and a *symbol* is an image that represents the object by convention.[35] Symbols are thus more abstract than indices, which are in turn more abstract than icons. Thus, in this study, we used 2 symbols: (a) large rectangular blocks and (b) filled ovals, along with an index: (c) circles with smile/frown faces inside them. The remaining 3 images would be classically defined as icons, as they were more anthropomorphic, that is, evocative of human form: (d) outlines of a person's head and shoulders, (e) male/female person icons similar to those used to denote restroom gender (which were matched to participants' gender), and (f) actual head-and-shoulder photographs of people of varied races (also matched to participants' gender). Dark blue icons denoted future cardiovascular events, and gray icons denoted nonevents (blue v. gray shirt color in the photo condition).

Our primary research questions involved whether icon type would influence risk recall, correlation of perceived risk with the presented risk statistics, or participants' ratings of the icon array as clear, helpful, and a preferable form of receiving such information.

Upon entering our survey, participants were given an introduction page that explained the purpose of the study, the anonymous nature of the research, and the expected time to take the survey. Participants also completed between 5 and 9 Web pages of survey materials for unrelated studies (cross-randomized across all 6 arms of this study) after completion of the primary and secondary measures for this study but before completion of individual difference measures (e.g., numeracy) and demographics. This design received institutional review board exempt status approval as anonymous survey research.

Measures and Covariates

We had 2 primary outcome measures. First, we measured recall of the provided risk percentage at the end of the survey (after demographics). Answers were coded as accurate if the respondent's answer was ±2% of the value. (Counting as accurate participants who were told their risk was ''more than 30%'' and who gave values larger than 32% did not qualitatively change our findings, so we report results using the more stringent accuracy criteria here.)

Second, we asked participants to answer 2 measures of their risk perceptions. Participants first rated ''How *big or small* does this risk *feel* to you?'' (emphasis in original) on a horizontal slider bar that recorded values from 0 (''extremely small'') to 100 (''extremely big''). They then rated ''How *likely* does it *feel* to you that you will actually get heart disease or stroke in the next 10 years?'' on another 0–100 point slider bar with end points labeled ''extremely unlikely'' and ''extremely likely.'' These questions were averaged to create an aggregate measure of risk perceptions (participants who only answered 1 of the 2 questions were included).

We also asked 3 graph evaluation items to evaluate user preferences about the different icon array displays. These questions asked respondents to use a 10-point Likert-type scale to rate how well the graphs described the risk of different side effects, how helpful the graphs were, and whether the participant would like to see risk information in this type of graph. Given prior research showing high correlations between these items,[10,11,36] we decided a priori to combine these ratings into a 3-item scale.

Because ample evidence exists that even highly educated adults can have poor numeracy skills (i.e., facility and comfort with quantitative health information such as risk statistics),[37–39] all study participants completed the Subjective Numeracy Scale (SNS),[40] a validated measure of quantitative ability and preference for receiving information in numerical form. The SNS has previously been shown to correlate

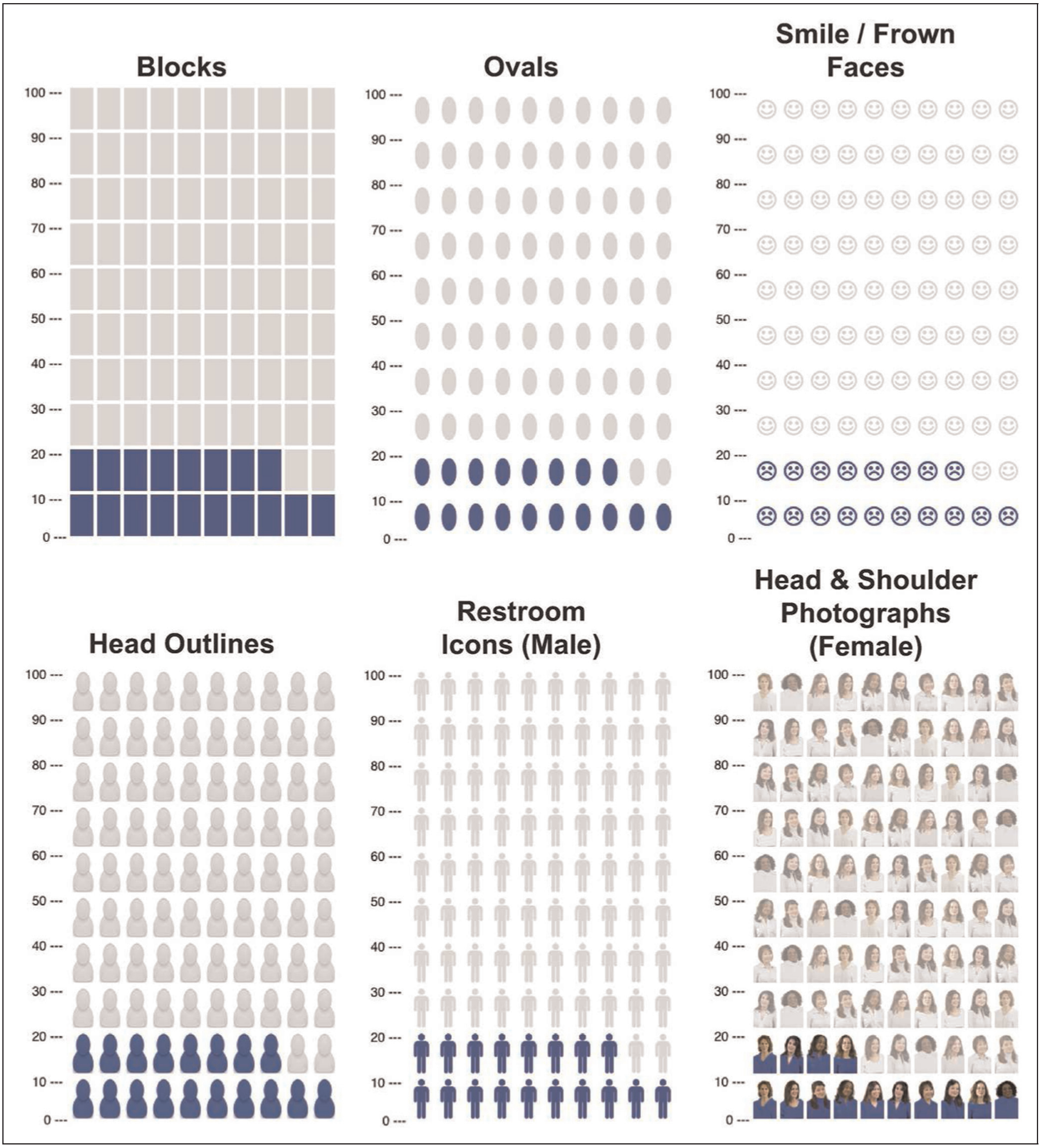

Figure 1 *The 6-icon array displays.*

with the ability to recall and comprehend both textual and graphical risk communications,[12,41] and recent work suggests that highly numerate people process icon arrays in different ways than do less numerate people.[21] A participant's SNS score is calculated as his or her mean rating across the 8 SNS questions and ranges from 1 (least numerate) to 6 (most numerate).

As recent research has suggested that visual displays are most useful to people with graphical literacy skills (a related, but distinct, construct from numeracy), we also measured graphical literacy using a reduced 6-question scale (to minimize respondent burden) comprised of items 5–6, 8–10, and 12 from Galesic and Garcia-Retamero's measure.[17,42] These questions were a representative sample of questions (selected based on predictive ability in our previous research) that presented readers with simple bar and line graphs and asked them to identify points, compare slopes, and draw inferences from these visuals. We counted the total number of correct answers as our measure of graphical literacy (range: 0–6).

Statistical Analyses

We first calculated mean rates of recall and graph preference scores for each experimental arm and then used logistic and linear regression models, respectively, with categorical variables for each arm to test whether icon choice resulted in significant differences in these outcomes.

To assess the relationship between perceived and reported actual risk, we also calculated the Pearson correlations between respondents' risk perception ratings and the risk estimates they received for each experimental condition. To assess whether the observed variations in these correlations across icon type conditions were significant, we conducted a linear regression analysis with (a) risk estimate as a continuous predictor and (b) icon type (as a set of categorical variables) and then (c) tested the significance of the risk estimate $\times$ icon type interaction terms (which are statistically equivalent to the correlations calculated above for each icon type). Since we did not have a priori hypotheses regarding which icon type would result in better versus worse performance, we adjusted the reference condition for each of the above regressions to be the experimental condition with the lowest accuracy rate, correlation, or mean preference score. Because including demographics in the models made little difference to the results (data not shown), we report only results from the simpler models.

In addition, we anticipated that not just the mean level of risk perception but also the pattern of perceived-to-actual risk correlations among different icon types might be different for participants with varying levels of numeracy and/or graphical literacy. To explore this possibility, we also performed subset analyses for each outcome variable by performing a median split by either numeracy or graphical literacy scores, grouping participants into lower versus higher subgroups, and rerunning both the correlations and the regression analyses for each numeracy or graphical literacy subgroup. All analyses were performed using STATA 12,[43] and all tests of significance were 2-sided and used alpha = 0.05.

RESULTS

Sample Description

In total, 1502 people between the ages of 35 and 74 reached the survey Web site, entered their information, and viewed their cardiovascular risk estimate in the icon array display. Mean estimated 10-year cardiovascular risk was 12% (median, 9%; interquartile range, 5%–18%). Characteristics for those participants who answered each demographic question are reported in Table 1. We observed a wide range of educational achievement, with 36.1% of participants having a bachelor's or higher college degree but also 22.6% having completed high school or less education. The SNS numeracy measure showed high reliability (Cronbach's alpha = 0.85), and mean SNS score was 4.14 (range, 1.0–6.0; s = 1.03). Mean score on our abbreviated graphical literacy scale was 3.95 (range, 1–6; s = 1.54). These 2 scales were correlated with each other (r = 0.41) but had substantial independent variance. In part because of this correlation, we present results for numeracy and graphical literacy separately below. Because questions about participant demographics came at the end of the survey, we do not know whether the demographics of those who dropped out after receiving their risk information differ from those who completed the survey.

Recall of the Risk Estimate

Overall, 75.4% of participants accurately recalled their 10-year cardiovascular risk estimate at the end of the survey, 19.1% provided overestimates (>2 percentage points), and 5.5% provided underestimates (>2 percentage points). However, the percentage with accurate recall varied significantly by arm, $\chi^2(5) = 15.84$, $P < 0.01$ (see Table 2), with corresponding reductions in the number of overestimators (the percentage of underestimates remained relatively constant). Comparing icon types in our logistic

Table 1 Participant Demographic Characteristics and Reported Cardiovascular Risk Estimates

Characteristic	Categories	Frequency (%)	$\bar{x}$ (s)
Age	35–39	136 (9.1%)	53.8 (9.7)
	40–49	366 (24.4%)	
	50–59	545 (36.3%)	
	60–69	380 (25.3%)	
	70–74	75 (5.0%)	
Gender	Male	687 (45.7%)	
	Female	815 (54.3%)	
Ethnicity	Hispanic (any race)	106 (8.2%)	
Race[a]	White	1030 (78.5%)	
	African American	205 (15.6%)	
	All others	91 (7.0%)	
Education	Less than high school	27 (2.1%)	
	High school only	269 (20.6%)	
	Some college/trade school	540 (36.0%)	
	Bachelor's degree	342 (26.1%)	
	Master's/doctorate	131 (10.0%)	
Subjective Numeracy Scale score	1.00–1.99	37 (2.8%)	4.14 (1.03)
	2.00–2.99	127 (9.7%)	
	3.00–3.99	353 (26.9%)	
	4.00–4.99	474 (41.3%)	
	5.00–5.99	307 (23.4%)	
	6.00	14 (1.1%)	
Graphical Literacy Scale score	0	27 (2.0%)	3.96 (1.54)
	1	64 (4.8%)	
	2	170 (12.8%)	
	3	203 (15.3%)	
	4	312 (23.5%)	
	5	315 (23.7%)	
	6	237 (17.9%)	
Reported 10-year cardiovascular risk	1%–4%	325 (21.6%)	11.99 (9.03)
	5%–9%	447 (29.8%)	
	10%–14%	274 (18.2%)	
	15%–19%	153 (10.2%)	
	20%–24%	82 (5.5%)	
	25%–30%	65 (4.3%)	
	30% to "more than 30%"	156 (10.4%)	

Note: The table reports results only for those respondents who completed each question or measure.
a. Respondents could mark more than 1 race.

regression analysis, we observed a clear pattern in which anthropomorphic icons (restroom icons, photographs, and head outlines) resulted in higher degrees of accuracy than more abstract icons such as blocks, which had the lowest recall rates (restroom v. blocks: odds ratio [OR] = 1.78, P = 0.01; photographs: OR = 1.79, P = 0.009; head outlines: OR = 1.47, P = 0.08). This pattern was particularly true for participants who scored below the median in graphical literacy or numeracy, although these respondents had generally lower accuracy rates than more numerate/graphically literate respondents regardless of icon type. Among respondents with higher graphical literacy or higher numeracy scores (who were consistently more accurate than respondents with lower scores, regardless of which icon was seen), only restroom icons resulted in significantly higher recall accuracy.

Risk Perceptions

Participants' ratings of the 2 risk perceptions questions were very similar in magnitude and were highly correlated (r = 0.71), so the aggregate measure of risk perceptions had high reliability (Cronbach's alpha = 0.83). Mean perceived CVD risk was 35.6 (s = 24.3)

Table 2 Percentage of Respondents Accurately Recalling Their 10-Year Cardiovascular Risk Estimate

		Graphical Literacy		Subjective Numeracy	
	Overall	Low	High	Low	High
Blocks	70.5	60.5	83.3	61.0	81.4
Ovals	71.3	64.2	83.1	61.5	79.6
Smile/frown faces	70.5	57.9	87.2	61.0	81.4
Head outlines	77.8	72.7[a]	85.7	77.2[a]	78.6
Restroom (M/F) icons	81.0[a]	71.6[a]	93.3[a]	73.9[a]	89.9[a]
Photographs (M/F)	81.0[a]	76.0[a]	87.0	76.4[a]	84.9

Note: All values are percentages.
a. $P < 0.05$ v. the condition with the lowest accuracy by logistic regression analysis.

Table 3 Correlation of Actual Displayed 10-Year Risk of Cardiovascular Events and Perceived Risk Reported by Participants

		Graphical Literacy		Subjective Numeracy	
	Overall	Low	High	Low	High
Blocks	0.24	0.31	0.33[a]	0.42[a,b]	0.19
Ovals	0.27	0.29	0.26[a]	0.25	0.33[a]
Smile/frown faces	0.14	0.16	0.22	0.15	0.27
Head outlines	0.09	0.15	–0.02	0.25	0.01
Restroom (M/F) icons	0.23	0.16	0.43[a]	0.20	0.39[a]
Photographs (M/F)	0.22	0.18	0.29[a]	0.07	0.35[a]

a. $P < 0.05$ v. the condition with the lowest correlation by regression analysis.
b. Because the overall test of icon type was not significant among low numerate participants, the significance of this pairwise comparison should be interpreted with caution.

on the 0–100 scale and did not differ significantly across experimental conditions. Overall, risk perceptions were moderately but significantly correlated with the actual risk data that had been presented ($r = 0.20$, $P < 0.001$). Per analysis of variance, this correlation did not vary significantly across icon types in the full sample (overall test of icon type × actual risk interaction when predicting perceived risk: $F = 1.33$, $P = 0.25$).

However, a more complicated and interesting pattern emerged once we divided our sample by graphical literacy scores and subjective numeracy scores (Table 3). Among highly graphically literate and highly numerate participants, icon type had a significant effect on the correlation between actual and perceived risk for cardiovascular disease (graphical literacy: $F = 2.75$, $P = 0.02$; numeracy: $F = 2.29$, $P = 0.04$). As shown in Table 3 (in the ''High'' columns), highly graphically literate and numerate participants who viewed icon arrays using restroom icons, photos, and ovals had notably higher correlations between perceived and actual risk than did similar participants who viewed other icon types. These differences in correlations were significant when compared with the group with the lowest correlations: those who saw head outlines.

By contrast, there was no overall effect of icon type on the correlation between actual and perceived risk among participants who scored below median on graphical literacy ($F = 0.76$, $P = 0.57$) and numeracy ($F = 1.89$, $P = 0.09$). Among less graphically literate/numerate participants, viewing restroom icons conferred no advantages vis-à-vis other icon types in terms of risk correlations. Thus, it appears from our data that the optimal icon type for increasing the association between perceived and actual risk depends on the individuals' numeracy and/or graphical literacy abilities.

Graph Preference Scores

As planned, we averaged participants' responses to the 3 graph preference questions to create an aggregate measure, which had high internal reliability (Cronbach's alpha = 0.91). Table 4 reports the mean graph preference scores for each of the 6 experimental

Table 4 Graph Preference Scores

		Graphical Literacy		Subjective Numeracy	
	Overall	Low	High	Low	High
Blocks	5.82	6.01	5.71	5.62	6.17
Ovals	6.13	6.10	6.43[a]	5.63	6.69
Smile/frown faces	5.83	5.67	6.36[a]	5.47	6.52
Head outlines	6.01	6.01	6.06	5.83	6.24
Restroom (M/F) icons	6.34[a]	6.33[a]	6.56[a]	6.04[a]	6.98[a]
Photographs (M/F)	6.31[a]	6.31[a]	6.44[a]	6.02[a]	6.63

Note: Scores are the mean of 3 questions, each on a 10-point (0–9) scale, where higher numbers represent greater preference.
a. $P < 0.05$ v. the condition with the lowest score by regression analysis.

conditions. Here, we observed significantly higher ratings for just 2 anthropomorphic icons: restroom icons and photographs (restroom: $\bar{x} = 6.34$, $s = 2.21$; photographs: $\bar{x} = 6.31$, $s = 2.11$). This pattern was consistently true for participants with lower numeracy and graphical literacy skills and was qualitatively similar (with somewhat smaller differences) for more numerate and graphically literate respondents. We note that the 2 preferred icon types were also the 2 conditions in which the icon was tailored by gender and that the effect of numeracy on graph preference scores appears substantially larger (regardless of icon type) than the effect of graphical literacy.

DISCUSSION

These results identify 2 important findings. First, icon type can matter. Although prior research conducted more than 15 years ago (and using very simplistic displays) suggested that risk perceptions were unaffected by icon type,[9] our work strongly suggests that the mental processing of risk information is indeed affected by the type of icons used in icon array displays. An icon array risk graphic that uses restroom icons (gender tailored to audience gender) appears to be easier to recall, more preferable, and (for more numerate and/or graphically literate people) more likely to result in well-calibrated risk perceptions than the exact same data shown in an array using blocks or ovals. Using actual head-and-shoulders photographs of a diverse group of people resulted in generally similar findings. However, the lack of any advantage to using photographs and the significant practical difficulties in creating and using them suggest that restroom icons are a more practical icon choice for many purposes.

Our second finding provides an important caveat regarding the benefits of using restroom icons. While restroom icons resulted in consistently good outcomes when used with more numerate and graphically literate people, the pattern was less clear for less numerate and graphically literate people. While restroom icons resulted in higher rates of recall and high graph preference ratings among less numerate and graphically literate participants, they conferred no advantage in terms of the correlations between perceived and actual risk. If anything, block icons (which performed particularly poorly among higher numerate/graphical literate participants) appeared to have somewhat higher correlations among less numerate/graphically literate participants than other icon types.

This finding is very consistent with recent work that has shown that less numerate people process icon arrays differently than more numerate people do. More numerate people tend to count icons, whereas less numerate people do not.[21] In our study, the restroom icons had the most ‘‘white space’’ (i.e., had larger gaps between the colored icons), thereby making them potentially more distinct and easy to count. By contrast, the blocks icons were solid filled rectangles with only a minimal amount of white border to allow each icon to be identifiable. Thus, it may have been harder visually to count blocks and easier to just roughly assess the relative area in the array that was colored blue versus the background gray color. We speculate that less numerate subjects found it easier to determine the ‘‘gist’’ (to use Reyna's term[44,45]) of the risk when blocks made the ‘‘part-whole’’ relationship visually simpler, whereas more numerate subjects found it easier whenever counting icons was facilitated. Further research is needed to separate the potential effects of icon shape versus visual density/spacing.

Our results are both similar to and different from those of a recent paper by Gaissmaier and others[32] that examined iconicity, that is, ‘‘how much a representation resembles what it is supposed to represent versus the extent to which it is an abstraction.''[32(p287)] In particular, we find relatively small differences between restroom icons and actual photographs, replicating the lack of difference that Gaissmaier and others observed between the upper levels of iconicity. The main contrast between our paper and theirs is our finding of substantial differences between highly abstract icons (blocks, ovals) and person-like icons (e.g., restroom icons), a finding that appears incongruent with their finding of limited differences between iconicity level 2 (abstract shapes) and 3 (stylized person icons) for several stimuli. There are several plausible hypotheses for this difference. First, in our study but not the Gaissmaier study, the person-like icons were gender-tailored to the participant, which may have increased attention or perception of the information as personally relevant. Second, our study reported actual personal risk of disease instead of a hypothetical treatment side effect, which may have amplified the effect of having the icon be anthropomorphic. Third, and perhaps most important, Gaissmaier and colleagues assessed knowledge as a primary outcome, whereas our primary outcome was subjective responses to the risk. It may be that icon type makes less of a difference in comprehension than in emotional response to the risk. Such explanations are clearly post hoc, however, and further research is needed to clarify the role of icon type and iconicity in patient reaction to icon-based risk displays.

Nonetheless, our findings have direct implications for the design of patient education materials used in both research and clinical practice. To start, our work suggests that developers of risk calculators and patient decision aids should carefully consider the choice of icon type in their risk graphics. The fact that icon type can affect important outcomes means that the decision about what icon to use has implications for both the effectiveness and usability of the resulting communication. As a result, our results suggest that researchers should clearly report in their publications exactly what icons are used in their risk displays. Additional research is needed both to confirm our findings and to identify the specific mechanisms that resulted in the differences we observed. In addition, clear guidance about what icons work best with different populations is needed for clinicians who may wish to use online tools (e.g., *iconarray.com*, a freely available icon array generator developed by several of the authors) to quickly create visual displays for use in clinical consultations.

Several limitations do temper the strength of our conclusions. First, although we asked participants to enter their actual health information and provided them with an estimate of their personal cardiovascular risk that is accurate for the information they provided, we cannot verify whether respondents entered accurate information into the calculator. Second, we recruited our sample from a commercial online survey panel. Thus, while it was diverse in both demographics and numerical/graphical abilities, it consisted exclusively of people who are willing to take surveys and who were computer-literate. Some respondents may have been more interested in completing a survey than in learning about their health risks, and thus it remains possible that our findings might not generalize to all patient contexts.

Nonetheless, this study reminds risk communicators of a finding we have seen over and over: Details matter. An icon array risk graphic that uses restroom icons is not interchangeable with an array showing the exact same data with smile faces. As communicators, we have an obligation not just to provide accurate information but also to match the formats used in our communications to patient needs.[46] Even if one might believe that icon type ''shouldn't matter,'' it appears that it does, and it behooves researchers and clinicians alike to adapt their methods to align with this evidence. Attention to details like these really does increase the likelihood that the decision aids, risk calculators, or education brochures we create will actually be able to guide and improve patients' medical decisions and outcomes.

REFERENCES

1. Hawley ST, Zikmund-Fisher BJ, Ubel PA, et al. The impact of the format of graphical presentation on health-related knowledge and treatment choices. Patient Educ Couns. 2008;73(3):448–55.

2. Tait AR, Voepel-Lewis T, Zikmund-Fisher BJ, et al. The effect of format on parents' understanding of the risks and benefits of clinical research: a comparison between text, tables, and graphics. J Health Commun. 2010;15(5):487–501.

3. Galesic M, Garcia-Retamero R, Gigerenzer G. Using icon arrays to communicate medical risks: overcoming low numeracy. Health Psychol. 2009;28(2):210–6.

4. Garcia-Retamero R, Galesic M. Who profits from visual aids: overcoming challenges in people's understanding of risks. Soc Sci Med. 2010;70(7):1019–25.

5. Fagerlin A, Wang C, Ubel PA. Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? Med Decis Making. 2005;25(4):398–405.

6. Garcia-Retamero R, Galesic M, Gigerenzer G. Do icon arrays help reduce denominator neglect? Med Decis Making. 2010;30(6):672–84.

7. Zikmund-Fisher BJ, Ubel PA, Smith DM, et al. Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. Patient Educ Couns. 2008; 73(2):209–14.

8. Waters EA, Weinstein ND, Colditz GA, et al. Reducing aversion to side effects in preventive medical treatment decisions. J Exp Psychol Appl. 2007;13(1):11–21.

9. Stone ER, Yates JF, Parker AM. Effects of numerical and graphical displays on professed risk-taking behavior. J Exp Psychol Appl. 1997;3(4):243–56.

10. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. Improving understanding of adjuvant therapy options by using simpler risk graphics. Cancer. 2008;113(12):3382–90.

11. Zikmund-Fisher BJ, Fagerlin A, Ubel PA. A demonstration of ''less can be more'' in risk graphics. Med Decis Making. 2010; 30(6):661–71.

12. Zikmund-Fisher BJ, Angott AM, Ubel PA. The benefits of discussing adjuvant therapies one at a time instead of all at once. Breast Cancer Res Treat. 2011;129(1):79–87.

13. Tait AR, Voepel-Lewis T, Zikmund-Fisher BJ, et al. Presenting research risks and benefits to parents: does format matter? Anesth Analg. 2010;111(3):718–23.

14. Feldman-Stewart D, Kocovski N, McConnell BA, et al. Perception of quantitative information for treatment decisions. Med Decis Making. 2000;20(2):228–38.

15. Feldman-Stewart D, Brundage MD, Zotov V. Further insight into the perception of quantitative information: judgments of gist in treatment decisions. Med Decis Making. 2007;27(1):34–43.

16. Garcia-Retamero R, Dhami MK. Pictures speak louder than numbers: on communicating medical risks to immigrants with limited non-native language proficiency. Health Expect. 2011;14 Suppl 1:46–57.

17. Okan Y, Garcia-Retamero R, Cokely ET, et al. Individual differences in graph literacy: overcoming denominator neglect in risk comprehension. J Behav Decis Mak. 2012;25(4):390–401.

18. Price M, Cameron R, Butow P. Communicating risk information: the influence of graphical display format on quantitative information perception—accuracy, comprehension and preferences. Patient Educ Couns. 2007;69(1–3):121–8.

19. Wright AJ, Whitwell SCL, Takeichi C, et al. The impact of numeracy on reactions to different graphic risk presentation formats: an experimental analogue study. Br J Health Psychol. 2009; 14(Pt 1):107–25.

20. McCaffery KJ, Dixon A, Hayen A, et al. The Influence of graphic display format on the interpretations of quantitative risk information among adults with lower education and literacy a randomized experimental study. Med Decis Making. 2012;32(4): 532–44.

21. Hess R, Visschers VHM, Siegrist M. Risk communication with pictographs: the role of numeracy and graph processing. Judgm Decis Mak. 2011;6(3):263–74. Available from: URL: http://journal.sjdm.org/11/10630/jdm10630.pdf

22. Carling CLL, Kristoffersen DT, Flottorp S, et al. The effect of alternative graphical displays used to present the benefits of antibiotics for sore throat on decisions about whether to seek treatment: a randomized trial. PLoS Med. 2009;6(8):e1000140.

23. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. BMJ. 2002;324:827–30.

24. Edwards A, Thomas R, Williams R, et al. Presenting risk information to people with diabetes: evaluating effects and preferences for different formats by a web-based randomised controlled trial. Patient Educ Couns. 2006;63(3):336–49.

25. Fortin JM, Hirota LK, Bond BE, et al. Identifying patient preferences for communicating risk estimates: a descriptive pilot study. BMC Med Inform Decis Mak. 2001;1:2.

26. Han PKJ, Klein WMP, Killam B, et al. Representing randomness in the communication of individualized cancer risk estimates: effects on cancer risk perceptions, worry, and subjective uncertainty about risk. Patient Educ Couns. 2012;86(1):106–13.

27. Brase GL. Pictorial representations in statistical reasoning. Appl Cogn Psychol. 2009;23(3):369–81.

28. Brown SM, Culver JO, Osann KE, et al. Health literacy, numeracy, and interpretation of graphical breast cancer risk estimates. Patient Educ Couns. 2011;83(1):92–8.

29. Ghosh K, Crawford BJ, Pruthi S, et al. Frequency format diagram and probability chart for breast cancer risk communication: a prospective, randomized trial. BMC Womens Health. 2008;8:18.

30. Kasper J, Heesen C, Köpke S, et al. Why not? Communicating stochastic information by use of unsorted frequency pictograms—a randomised controlled trial. Psychosoc Med. 2011;8:Doc08.

31. Wong ST, Pérez-Stable EJ, Kim SE, et al. Using visual displays to communicate risk of cancer to women from diverse race/ethnic backgrounds. Patient Educ Couns. 2012;87(3):327–35.

32. Gaissmaier W, Wegwarth O, Skopec D, et al. Numbers can be worth a thousand pictures: individual differences in understanding graphical and numerical representations of health-related information. Health Psychol. 2012;31(3):286–96.

33. D'Agostino RB Sr, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117(6):743–53.

34. Witteman HO, Zikmund-Fisher BJ, Waters EA, et al. Risk estimates from an online risk calculator are more believable and recalled better when expressed as integers. J Med Internet Res. 2011;13(3):e54.

35. Mullet K, Sano D. Designing Visual Interfaces: Communication Oriented Techniques. Englewood Cliffs (NJ): SunSoft Press; 1995.

36. Zikmund-Fisher BJ, Witteman HO, Fuhrel-Forbis A, et al. Animated graphics for comparing two risks: a cautionary tale. J Med Internet Res. 2012;14(4):e106.

37. Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. Med Decis Making. 2001;21(1):37–44.

38. Nelson W, Reyna VF, Fagerlin A, et al. Clinical implications of numeracy: theory and practice. Ann Behav Med. 2008;35(3): 261–74.

39. Peters E, Hibbard J, Slovic P, et al. Numeracy skill and the communication, comprehension, and use of risk-benefit information. Health Aff. 2007;26(3):741–8.

40. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, et al. Measuring numeracy without a math test: development of the subjective numeracy scale (SNS). Med Decis Making. 2007;27(5):672–80.

41. Zikmund-Fisher BJ, Smith DM, Ubel PA, et al. Validation of the subjective numeracy scale (SNS): effects of low numeracy on comprehension of risk communications and utility elicitations. Med Decis Making. 2007;27(5):663–71.

42. Galesic M, Garcia-Retamero R. Graph literacy: a cross-cultural comparison. Med Decis Making. 2011;31(3):444–57.

43. Stata Statistical Software, version 12. College Station (TX): Stata Corporation; 2011.

44. Reyna VF. How people make decisions that involve risk: a dual-processes approach. Curr Dir Psychol Sci. 2004;13(2):60–6.

45. Reyna VF, Nelson WL, Han PK, et al. How numeracy influences risk comprehension and medical decision making. Psychol Bull. 2009;135(6):943–73.

46. Zikmund-Fisher BJ. The right tool is what they need, not what we have: a taxonomy of appropriate levels of precision in patient risk communication. Med Care Res Rev. 2013;70(1 Suppl):37S–49S.