ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Organizational Behavior and Human Decision Processes 90 (2003) 19–36

ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES

www.elsevier.com/locate/obhdp

Foreground:background salience: Explaining the effects of graphical displays on risk avoidance

Eric R. Stone,[a,*] Winston R. Sieck,[b] Benita E. Bull,[a] J. Frank Yates,[b] Stephanie C. Parks,[a] and Carolyn J. Rush[a]

[a] *Wake Forest University, Winston-Salem, NC 27109, USA*
[b] *University of Michigan, MI, USA*

Abstract

The purpose of this research was to determine the mechanisms underlying the graphical effect identified by Stone, Yates, and Parker (1997), in which graphical formats for conveying risk information are more effective than numerical formats for increasing risk-avoidant behavior. Two experiments tested whether this graphical effect occurred because the graphical formats used by Stone et al. highlighted the number of people harmed by the focal hazard, causing the decisions to be based mainly on the number of people harmed (which we label the "foreground") at the expense of the total number of people at risk of harm (which we call the "background"). Specifically, two graphical formats were developed that displayed pictorially both the number of people harmed and the total number at risk, and use of these display formats eliminated the graphical effect. We thus propose that the previously discussed graphical effect was in fact a manifestation of a more general foreground:background salience effect, whereby displays that highlight the number of people harmed at the expense of the total number of people at risk of harm lead to greater risk avoidance. Theoretical and practical implications are discussed.

1. Introduction

The issue of how best to communicate risk information has become an increasingly important part of the risk assessment discipline (see, e.g., Fischhoff, 1995; Fisher, 1991). Although the risk communication field is very broad, most of the work has focused on one of three goals: (a) increasing knowledge about the risks, (b) modifying risk-relevant behavior, and (c) facilitating cooperative conflict resolution (Lipkus & Hollands, 2000; Rohrmann, 1992). The focus of the present work is on the second of these goals, modifying risk-relevant behavior. Consider, for example, communication attempts designed to influence teenagers not to take drugs or to practice safe sex. Gaining an understanding of what techniques are the most effective for these sorts of goals can assist these intervention attempts in a number of situations. More specifically, the present work examined methods for modifying

* Corresponding author. Fax: +1-336-758-4733.
E-mail address: estone@wfu.edu (E.R. Stone).

0749-5978/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0749-5978(03)00003-7

risk-relevant behavior by means of alternative methods of displaying risk magnitude information.

Much of the research on modifying risk-relevant behavior has examined situations where the probability of the undesired event is quite small. Indeed, a large number of the hazards that have led to intervention attempts (e.g., of dying from an overdose of a drug or contracting AIDS from practicing unsafe sex) have low probabilities of occurring per instance. However, when these actions accumulate over time and people, the resulting effects become quite serious. In addition, most people have limited practical experience dealing with low-probability events. A number of researchers have thus suggested that people have difficulty reasoning on the basis of these low-probability risks (see, e.g., Camerer & Kunreuther, 1989; Covello, von Winterfeldt, & Slovic, 1986; Fisher, 1991; Fisher, McClelland, & Schulze, 1989; Halpern, Blackman, & Salzman, 1989; Lipkus & Hollands, 2000; Magat, Viscusi, & Huber, 1987; Roth, Morgan, Fischhoff, Lave, & Bostrom, 1990; Stone, Yates, & Parker, 1994, 1997). For example, Fisher et al. (1989) suggested that people either dismiss low-probability events entirely or else focus primarily on the size of the expected loss (similar arguments are given by Halpern et al. (1989), Magat et al. (1987), and Stone et al. (1994)). This type of argument is in keeping with prospect theory (Kahneman & Tversky, 1979), which suggests that small probabilities are either "edited" to zero or else overweighted, as well as with fuzzy trace theory (Reyna & Brainerd, 1991), which suggests that people encode the gist of the available information (e.g., that the probabilities are quite small) and, whenever possible, reason on the basis of this gist rather than on finer distinctions among the probabilities.

These arguments have motivated a variety of different proposed techniques for conveying low-probability risk magnitudes. In one classic study, Slovic, Fischhoff, and Lichtenstein (1978) told participants either that the chance of experiencing at least one disabling injury when driving without a seatbelt is .00001 for each trip or that the probability is .33 over 50 years of driving. Although these statistics are formally equivalent, the latter frame, aggregated over 50 years of driving, produced more risk-avoidant behavior on the part of the participants. More recently, Stone et al. (1994) showed that presenting risk information in relative risk form (e.g., that a safer product reduces the risk to half that of another product) led to more risk-avoidant behavior than simply giving the risk magnitudes for the two products (see also Baron, 1997; Halpern et al., 1989). Siegrist (1997) showed that, under certain conditions, providing risk information via a frequency format (e.g., 600 out of 1,000,000 people will die) leads to more risk-avoidant behavior than simply providing the risk magnitudes in probability (incidence rate) form. Weinstein, Kolb, and Goldstein (1996) used the time intervals between expected events to communicate risk magnitudes, and showed that as long as the time intervals were long, presenting them in addition to the risk magnitudes led to a smaller perceived need for action (less risk avoidance).

A number of authors (e.g., Covello, Sandman, & Slovic, 1988; Keeney & von Winterfeldt, 1986) have proposed that the presentation of risk information in graphical form should be an especially effective means of increasing risk-avoidant behavior. Stone et al. (1997) demonstrated empirically the efficacy of such recommendations, by showing that graphical techniques for displaying risk information can indeed be more effective than simply providing numerical information for highlighting the risk reduction accorded by a safer product. Specifically, they provided information about the risk associated with either tire blowouts or the development of periodontal disease. For example, they told participants that with "Standard Toothpaste," 30 out of 5000 people would develop periodontal disease. They then informed participants about "Improved Toothpaste," which is identical to Standard Toothpaste except for the fact that it reduces the risk to 15 out of 5000.

In their first experiment, the number of people developing periodontal disease was displayed either numerically by the numbers "30" and "15" or graphically by means of stick figures illustrating the people developing periodontal disease. Participants were willing to pay more for the safer product when the risk was displayed via the stick figures format than by the numbers display. In two subsequent experiments, Stone et al. demonstrated that using stick figures to represent the risk was not essential to the effect, in that similar effects held for other graphical displays, in particular, asterisks and bar graphs. (See Fig. 1 for the numbers and asterisks formats used in the Stone et al. study.)

Although Stone et al. showed that depicting risk information graphically as opposed to numerically is a useful technique for increasing risk-avoidant behavior, they only speculated as to the mechanisms responsible for this "graphical effect." As Lipkus and Hollands (2000) discuss, this somewhat atheoretical approach to the study of graphical displays of risk is common, despite the fact that theoretical explanations for these sorts of effects are necessary both to integrate the findings on communicating risk information as well as to determine which types of graphical formats would be most effective in specific situations. The work of Stone et al. (1997) showed that for the particular modes of presentation they used, the choice of graphical format did not matter. However, it is important to understand precisely the mechanism underlying the graphical effect, so that any boundary conditions on this effect can be anticipated.

The goal of the present research was to gain a deeper understanding of the reasons behind the graphical effect identified by Stone et al. Consider again the task

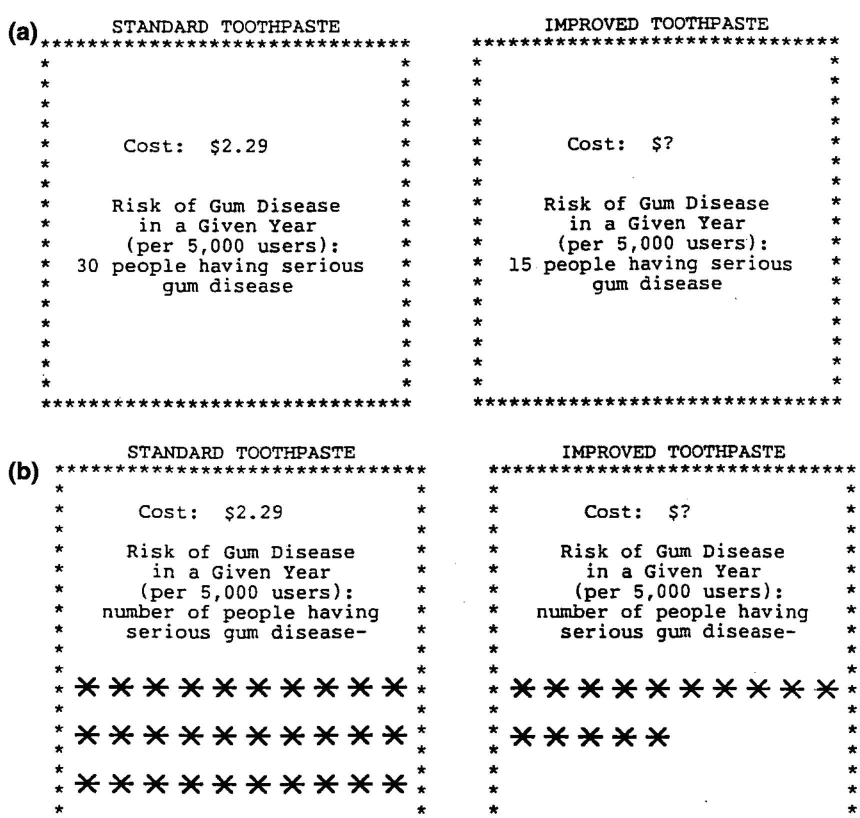

Fig. 1. Two examples of display formats used in the study by Stone et al. (1997): (a) the numbers format; (b) the asterisks format.

confronting the participants in Stone et al.'s study. For both the "Standard" and "Improved" products, they were given risk information in the form of a ratio, that x out of y would develop periodontal disease. Using the terminology introduced by Halpern et al. (1989), we refer to the number of people developing periodontal disease (x) as the "foreground," and the number of people at risk (y) as the "background." For example, for the toothpaste product used by Stone et al., the background was 5000 for both the Standard and Improved products, and the foreground was 30 for the Standard product and 15 for the Improved product. In this and the other scenarios used by Stone et al., the information in the foreground (30 vs. 15) suggests a fairly strong risk reduction, but the information in the background suggests that the risk reduction is not so strong (since the number of people at risk of harm is quite large, making the resultant incidence rates relatively small). To the extent that attention is drawn to one of these components more than to the other, then, it should be possible to make the risk reduction appear either large or small.

Thus, we suggest that the "graphical effect" discussed by Stone et al. was in fact a manifestation of a more general "foreground:background salience effect," in that the graphical conditions used by Stone et al. were effective for increasing risk avoidance because they highlighted the number of people harmed at the expense of the number of people at risk of harm. More specifically, we are hypothesizing that in their numbers condition, participants reason based on relatively equal consideration of the foreground and background information. Given that the ratio of the foreground to the background is small, participants see the initial risk as being "low." They thus perceive the improvement with the safer product as being relatively small, and offer to pay only a small amount more for the safer product. Fig. 2a illustrates the extreme case where the foreground and background receive perfectly equivalent consideration, though it is possible that in actuality the foreground and background do not receive exactly equal weight when participants are presented with a numerical display.

In the graphical conditions used by Stone et al., however, we propose that participants' attention is drawn immediately to the foreground information. This hypothesis is based on the assumption that the graphical display of the foreground information is more salient than the numerical description of the background information, and that, in accord with the conclusions reached by Sanfrey and Hastie (1998), information that is particularly salient will have the strongest effect on any decision based on that display (see also Jarvenpaa, 1990). Thus, participants' representations of the decision situation will be dominated by the foreground information, leading them to perceive the original risk as being "moderately large" and the reduction in risk as significant, which leads them to be willing to pay a moderate amount more for the safer product. Fig. 2b illustrates the extreme case where attention is devoted solely to the foreground.

To the extent that this mechanism is correct, it extends the work of Stone et al. in two ways. First, it suggests that the previously documented effectiveness of graphical displays occurs due to their effect on the perceived risk reduction. It is worth emphasizing that there are other potential mechanisms for the graphical effect that do not imply that the risk reduction is perceived as being greater when graphical versus numerical displays are used. As Stone et al. (1997) discussed, the extent of the risk reduction is only one of many considerations that would determine how much a person would be willing to pay for a safer product, and some of these considerations will favor low prices for the safer product, such as the equivalence of the two toothpastes on other factors and additional uses for which the money could be spent. It is plausible that, by calling attention to the risk reduction, graphical displays increase the importance associated with the risk reduction in relation to these other considerations. If that is the case, then participants should see the risk reduction as

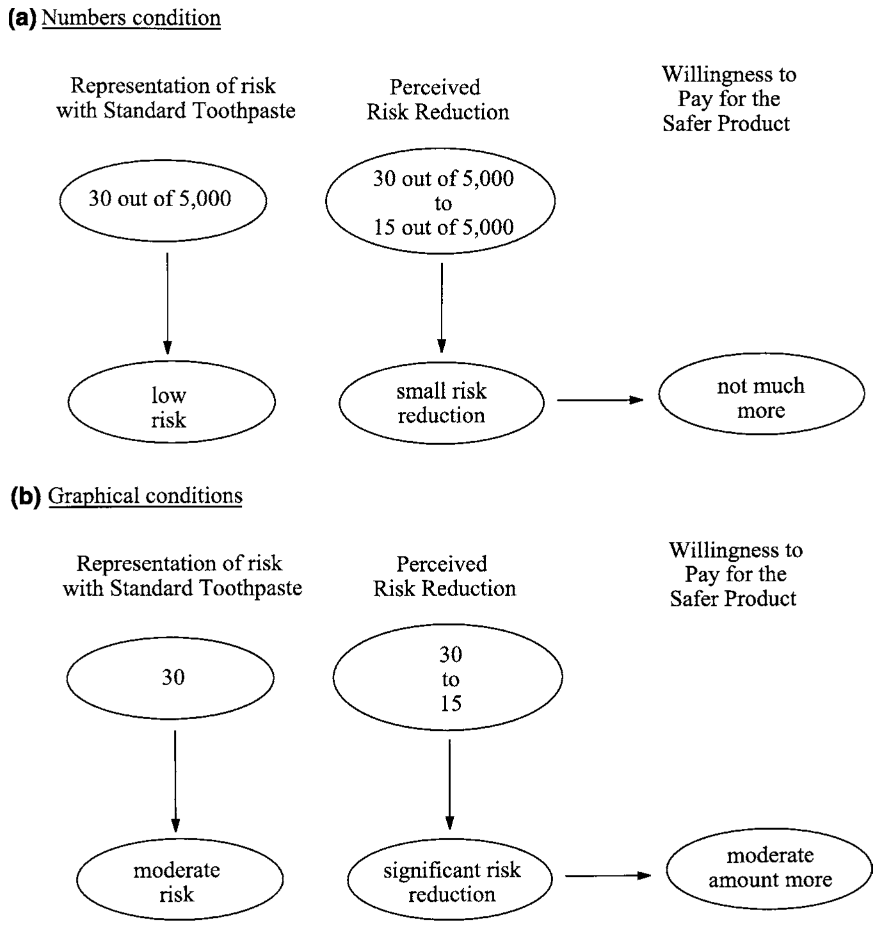

Fig. 2. Proposed mechanism for the difference in risk avoidance between the numbers and graphical conditions in the study by Stone et al. (1997): (a) the numbers condition; (b) the graphical conditions.

being more important with the graphical display, but not see the risk reduction itself as being any larger.

Second, our proposed foreground:background salience explanation suggests that the previously named graphical effect is not driven by a graphical mode of presentation per se, but instead by the particular aspect of the graphical formats used by Stone et al., whereby the foreground, but not the background, information was highlighted via the graphical mode of presentation. In other words, this explanation suggests that the greater effectiveness of graphical formats will not hold for all graphical modes of presentation. The logic behind the present research was to modify the graphical formats so that they did not highlight the foreground information to a greater extent than the background information, and determine whether or not the effectiveness of graphical formats in inducing risk avoidance remained after that change.

2. Experiment 1

The purpose of our first study was to test the previously discussed implications of our proposed foreground:background salience mechanism for the graphical effect documented by Stone et al. First, we needed to construct a graphical presentation mode that emphasized the background information to the same extent as the foreground information. One graphical format that meets these requirements is a pie chart. This mode of presentation does display the risk level graphically, but, unlike the graphical formats used by Stone et al., displays both the foreground and the background risk information graphically. Thus, to the extent that the graphical effect occurred because the graphical formats highlighted the foreground information at the expense of the background information, using pie charts to display the risk information should eliminate the graphical effect.

Additionally, we designed this experiment to test whether any effect of display format on risk avoidance occurs because of a difference in perceived size of the risk reduction. If this explanation is correct, then those display formats that are most successful at increasing the perceived size of the risk reduction should have a corresponding effect on the level of risk avoidance. To test this possibility, we included questions designed to measure the perceived risk reduction as well as the level of risk avoidance.

2.1. Method

Participants. Participants were 269 students from Wake Forest University and the University of Michigan, who participated as partial fulfillment of a psychology course requirement. One participant was excluded from the analysis for failure to respond to the entire questionnaire.

Materials. Our materials were adapted from those used by Viscusi and Magat (1987), whose procedures were designed to evaluate consumers' tradeoffs between a product's risk and cost. As part of their study, Viscusi and Magat presented participants with two hypothetical bleaches, each with an associated level of risk. Participants were then given the price of the riskier bleach, and asked how much they would be willing to pay for the safer one. This technique allows one to examine directly participants' professed behavior as a result of varying the risk display format.

As discussed elsewhere (Stone et al., 1994; Stone et al., 1997), we employed essentially the same procedure as Viscusi and Magat, except that we used products likely to be familiar to undergraduates, in particular, with risk information concerning the possibility of contracting serious gum disease (Brown, Oliver, & Loe, 1990). Specifically, we told participants there was a certain risk of developing periodontal disease if one used what we called "Standard Toothpaste," and that the manufacturer was considering marketing a new brand of toothpaste. This brand would be identical in all respects to the former product, except that it would reduce the risk by a given amount. The participant was then given the price of the Standard brand of toothpaste, and asked how much she would be willing to pay for the safer product ("Improved Toothpaste"). Finally, the risk information was presented to the participant in one of three formats.

The first two formats, numbers and asterisks, were the same as those used in the Stone et al. (1997) study (see Fig. 1). Specifically, the numbers format provided the risk information for both the Standard and Improved products in terms of the number of people out of 5000 who would be expected to develop periodontal disease. The asterisks format was identical to the numbers display, except that instead of being given a number, participants were shown asterisks representing the number of individuals who would develop periodontal disease. Finally, a pie chart format presented the same information that was provided in the previous two displays, but did so by means of pie charts, for reasons discussed previously. Since that format required more space than was used for the display formats in the Stone et al. study, the size of the boxes for the numbers and asterisks formats was increased compared

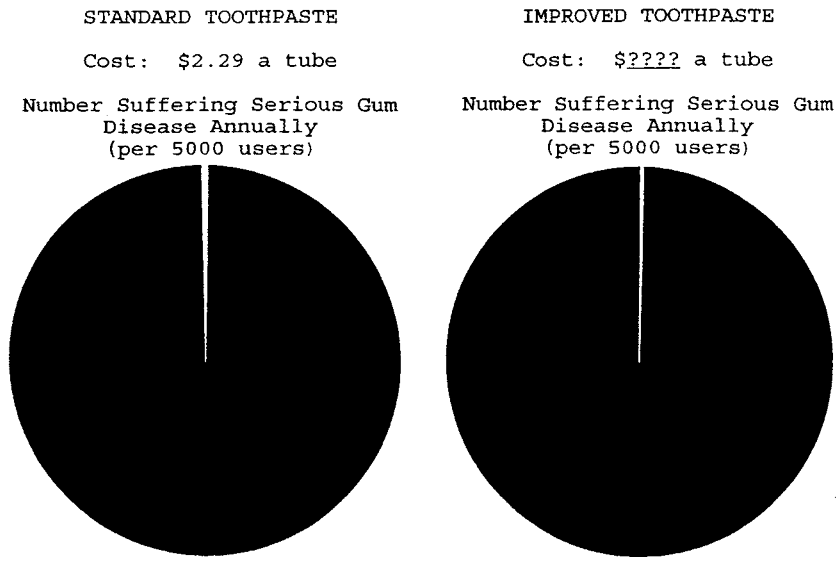

Fig. 3. Display of the pie charts format used in Experiment 1.

to what was used in that study as well. See Fig. 3 for the pie charts format used in this study.

To address the hypothesis that the graphical format was effective by means of increasing the perceived risk reduction, additional questions were developed using 7-point Likert-type scales. Although we asked some other questions for exploratory purposes, the two questions that were most relevant asked about the size of the perceived risk reduction and the significance of the risk reduction. In particular, the "Risk Reduction Size" and the "Risk Reduction Significance" questions were, respectively

> Risk Reduction Size = "Do you believe the decrease in the risk resulting from the IMPROVED TOOTHPASTE was (circle one of the choices)": (1 = none, 7 = incredibly big)
>
> Risk Reduction Significance = "Please indicate how significant you believe the reduction of the risk was when going from the STANDARD toothpaste to the IMPROVED toothpaste by circling a number on the scale below. *In other words:* 'When I compare the amount of risk associated with the IMPROVED toothpaste to the risk associated with the STANDARD toothpaste, I'd say the difference is...' (circle a # between 1 and 7 on the scale below)" (1 = insignificant, 7 = highly significant)

Note that differences in risk reduction significance, like differences in risk avoidance, could occur for reasons besides a perceived difference in risk reduction size. For example, the graphical format could make the risk reduction appear more significant by making periodontal disease appear to be more serious. Thus, if the display format had a greater effect on the risk reduction significance question than on the risk reduction size question, this would provide evidence that the graphical effect was operating, at least in part, by some mechanism other than by increasing the size of the perceived risk reduction. If, however, responses were similar to these two items, then this would provide further evidence that the graphical effect was driven primarily by an increase in the perceived size of the risk reduction.

Procedure. Each participant was presented with the toothpaste scenario with the risk depicted in one of the three formats. At the bottom of the page the participants were instructed to indicate how much they would be willing to pay for the improved toothpaste. On a second page the participants responded to the additional questions addressing the risk reduction.

2.2. Results

Although we report the means and standard deviations of the untransformed scores, we used a logarithmic transformation of the actual prices given by the participants for all inferential tests and correlations to reduce the impact of the positive skew in the original data (see Howell, 1992, p. 311). The means were compared via pairwise comparisons. All comparisons between the asterisks and numbers conditions were conducted via one-tailed t tests, but the comparisons involving pie charts were done via two-tailed t tests to account for the novel nature of that condition.

Table 1 presents the means and standard deviations per display format for the prices participants were willing to pay for the safer product as well as for the risk reduction questions. The difference found by Stone et al. between the asterisks and the numbers condition was replicated in this experiment, although the difference only reached marginal significance. Specifically, participants were willing to pay $.21 more for the safer toothpaste in the asterisks than in the numbers condition, $t(265) = 1.35$, $p = .09$. As predicted, however, highlighting the background information by means of pie charts eliminated the graphical effect. In particular, participants were willing to pay $.61 more to reduce the risk when given risk information by means of asterisks than by means of pie charts, $t(265) = 4.96$, $p < .001$. In fact, using pie charts rather than asterisks reversed the graphical effect, as participants were willing to pay $.41 more when presented with numbers than with pie charts, $t(265) = 3.64$, $p < .001$.

Participants responded similarly to the risk reduction size and risk reduction significance questions. In particular, the correlation between those two items was .76. Moreover, the correlations between each of those questions and the amount participants were willing to spend were nearly identical (.36 for risk reduction size and .35 for risk reduction significance), and were significantly different from 0, both p's $< .001$.

As shown in Table 1, for both risk reduction questions, participants gave the highest risk reduction estimates in the asterisks condition, followed by the numbers condition, followed by the pie charts condition, in accord with the findings for the amount paid variable. In both cases, the difference between the pie charts and the other formats was highly significant (all p's $< .01$), although the difference between the asterisks and numbers conditions reached significance only for the risk reduction significance question, $t(265) = 2.05$, $p = .02$.

Table 1
Means of dependent measures per condition in Experiment 1

Dependent measure	Condition		
	Numbers	Asterisks	Pie charts
Average amount paid [Standard = $2.29]	$3.39 ($0.81)	$3.60 ($1.18)	$2.99 ($0.59)
Risk reduction size	3.93 (1.30)	4.07 (1.12)	2.97 (1.05)
Risk reduction significance	3.37 (1.30)	3.73 (1.11)	2.87 (1.09)

Note. Numbers in parentheses are standard deviations. Sample sizes varied per condition, but ranged from $n = 88$–90. Risk reduction scores ranged from 1 to 7, and larger numbers represent a greater perceived risk reduction.

2.3. Discussion

These results support the idea that the graphical effect discussed by Stone et al. (1997) occurred because the graphical formats used in their experiment highlighted the foreground information at the expense of the background information. Indeed, using pie charts not only eliminated the graphical effect, but actually reversed it. We discuss potential reasons for this reversal in Section 4.

It also appears that the greater effectiveness of certain display formats results primarily from emphasizing the extent of the risk reduction. Participants responded similarly to the risk reduction size and significance questions, and both these items were affected by the format type in the same manner as was the price participants were willing to pay for the safer product. The only result inconsistent with this explanation is the failure of the numbers–asterisks difference to reach significance for the risk reduction size question. Given the similarity of the other findings with the two risk reduction questions and the number of analyses conducted, however, it seems plausible that this non-significance simply reflects a Type 2 error.

One potential limitation of Experiment 1 is that it was difficult to tell from the pie charts exactly how many people would have developed periodontal disease in the two conditions. On the one hand, this is a direct reflection of the fact that the background in addition to the foreground information was displayed graphically, making the resultant slice of the pie chart representing the foreground quite small. Nonetheless, with the other display formats, the foreground information was either stated explicitly (in the numbers condition), or easily calculable (by counting the number of asterisks.) It could be the case, then, that people perceived the risk statistics less accurately with the pie charts compared to the other display formats. The purpose of Experiment 2 was to address this concern, as well as to determine whether the effect found with pie charts would generalize to an additional display format that highlighted both the foreground and the background information.

3. Experiment 2

The basic approach taken in Experiment 2 was the same as in the first study. However, to provide tighter control, the asterisks condition was replaced with a bar graph condition, similar to the one used in the Stone et al. study, and the pie charts condition was replaced by a stacked bar chart, which displayed both the foreground and the background information graphically (see Fig. 4). This change accomplished two goals. First, it allowed us to test the generalizability of the results found in Experiment 1. If the same results are found with two different formats that display both the foreground and background information graphically, that would provide strong support for our foreground:background salience account of Stone et al.'s graphical effect. Second, using bar graphs made it possible to display the relevant risk information numerically on the graphs. This allowed us to address the concern that perhaps the participants incorrectly judged the actual risk facts contained in the pie charts condition.

3.1. Method

Participants. Participants were 414 students from Wake Forest University and the University of Michigan, who participated in the experiment as partial fulfillment of a psychology course requirement.

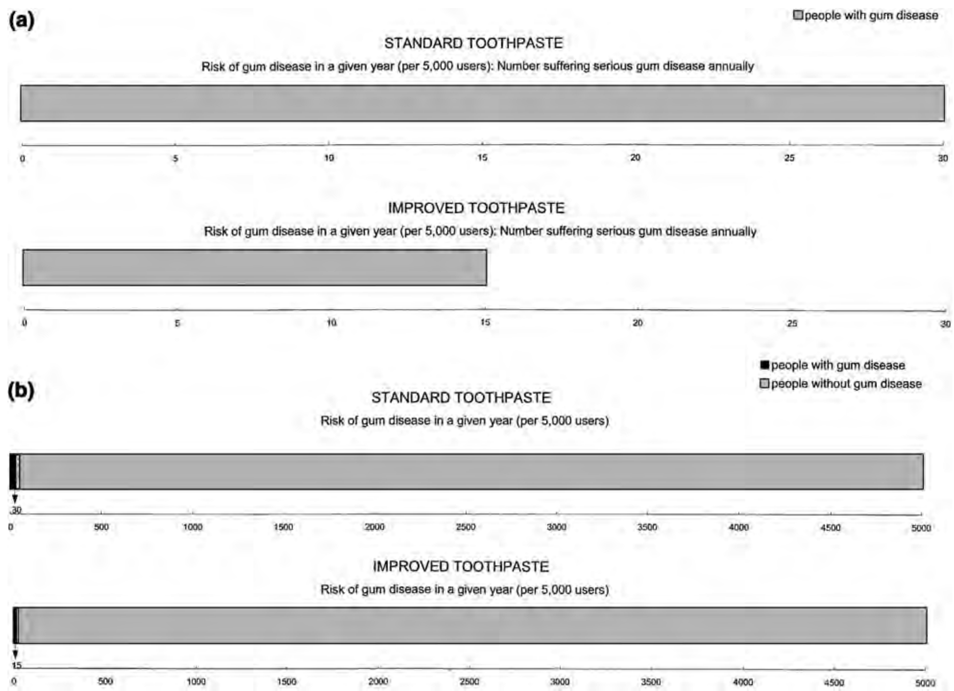

Fig. 4. Display of the graphical conditions used in Experiment 2: (a) the bar graphs format; (b) the stacked bar graphs format.

Materials. As in Experiment 1, participants were presented with two hypothetical brands of toothpaste, Standard and Improved, each with an associated risk of developing gum disease. The information provided was identical to that given in the first experiment, except that this study replaced the asterisks format with bar graphs and the pie charts format with stacked bar graphs. Specifically, the numbers format was the same as that used in Experiment 1, and presented the numbers 30 and 15 under Standard Toothpaste and Improved Toothpaste, respectively. The bar graph display used bars to represent the 30 and 15 people who would develop gum disease (see Fig. 4a). The stacked bar graph format visually presented the ratio of the number of people harmed by gum disease to the total number of people at risk of harm by displaying the number of people who developed gum disease (30 or 15) as a section within a bar graph representing the total number of people at risk (5000) (see Fig. 4b). Finally, to make these displays easier to read, they were printed on 8.5” by 14” paper rather than on 8.5” by 11” paper, and via a landscape layout rather than a portrait layout.

Also as in Experiment 1, participants responded to a number of additional questions designed to increase our understanding of why they made the choices that they did. Two of these were the “Risk Reduction Size” and “Risk Reduction Significance” questions used previously. In addition, in this study we asked participants to describe more precisely why they made the choices that they did. Specifically, participants were provided with potential considerations that might have influenced their decisions regarding how much they were willing to pay for Improved Toothpaste, and were asked to rate how much each affected their decision on a scale of 1 (had no impact) to 7 (had a major impact). Although we asked a number of these questions for exploratory purposes, the two most relevant to our present purposes were: “I examined the difference between the number of people who would develop periodontal disease (30 vs. 15)” (foreground comparison), and “I examined the difference in the chance of developing periodontal

Table 2
Mean willingness to pay by condition in Experiment 2

	Condition		
	Numbers	Bar graphs	Stacked Bar graphs
Average amount paid [Standard = $2.29]	$3.23 ($0.68)	$3.94 ($2.11)	$3.12 ($0.73)
Percentage increase from standard price	41%	72%	36%

Note. Numbers in parentheses are standard deviations. Sample sizes varied per condition, but ranged from $n = 127$–144.

disease between the two toothpastes (30/5000 vs. 15/5000)" (foreground:background comparison).

Procedure. Each participant was randomly assigned to receive the risk information for the two toothpastes in either the numbers, bar graph, or stacked bar graph format. Participants were told that the Standard Toothpaste sold for $2.29, and were asked to indicate how much they would be willing to pay for Improved Toothpaste, given the reduction in the risk of gum disease. Then, they responded to the additional questions on a separate page.

3.2. Results

A number of participants failed to respond to various portions of the questionnaire. In particular, three participants did not answer the question about how much they were willing to pay, 12 participants did not respond to the perceived risk reduction questions, and 24 participants did not respond to the two questions described above about how they went about making the decision. In addition, data from two participants were adjusted because their responses were deemed to be not feasible. Specifically, these participants indicated that they would pay less for Improved Toothpaste than for Standard Toothpaste, and we thus adjusted their responses to $2.29, the price of Standard Toothpaste.[1] After these adjustments were made, we used a logarithmic transformation of these prices for all inferential tests to account for the positive skew in the data, although as with Experiment 1, we report the descriptive statistics on the untransformed scores. Also as in Experiment 1, all comparisons between the conditions used in the Stone et al. study (in this case, numbers and bar graphs) were conducted via one-tailed pairwise comparisons, but the tests involving the new condition (in this case, stacked bar graphs) were done via two-tailed pairwise comparisons. The one exception to this general rule involved the participants' ratings of their decision considerations, where we conducted two-tailed tests throughout since we felt there was less of a firm basis for any predictions.

Table 2 presents the mean prices participants in each condition were willing to pay for Improved Toothpaste. A priori pairwise comparisons revealed that, overall, participants in the bar graph condition were willing to pay a higher price for Improved Toothpaste than were participants in both the numbers condition, $t(408) = 4.55$, $p < .001$, and the stacked bar graph condition, $t(408) = 5.75$, $p < .001$. Specifically, participants who received the information in the bar graph format were willing to pay $0.71 more for Improved Toothpaste than were participants who received the information in the numbers format, and $0.83 more than

[1] Specifically, one participant stated that he or she would pay $.50 for the Improved Toothpaste, and the other participant stated a willingness to pay $2.20 for the Improved Toothpaste. We analyzed the data leaving these data intact as well, and this produced no qualitative differences in the results from what we found when changing these scores to $2.29.

Table 3
Mean risk reduction ratings by condition in Experiment 2

Risk reduction measure	Condition		
	Numbers	Bar graphs	Stacked bar graphs
Risk reduction size	3.90 (1.26)	4.45 (1.33)	3.27 (1.24)
Risk reduction significance	3.41 (1.30)	3.96 (1.32)	2.88 (1.26)

Note. Numbers in parentheses are standard deviations. Sample sizes varied per condition, but ranged from $n = 125$–140. Risk reduction scores ranged from 1 to 7, and larger numbers represent a greater perceived risk reduction.

were participants who received the information in the stacked bar graph format. Participants in the numbers condition were willing to pay $.12 more on average for the Improved Toothpaste than were participants in the stacked bar graph condition, but this difference did not reach significance, $t(408) = 1.27$, $p = .21$.

The mean responses per condition to the questions addressing the perceived size and significance of the risk reduction are given in Table 3. For both these questions, the mean response given by participants in the bar graph condition was greater than the mean response given by participants in the numbers condition, which was greater than the mean response given by participants in the stacked bar graph condition. Specifically, the size of the risk reduction resulting from Improved Toothpaste was perceived to be greater in the bar graph condition than in both the numbers condition, $t(399) = 3.52$, $p < .001$, and the stacked bar graph condition, $t(399) = 7.50$, $p < .001$. Additionally, participants in the numbers condition perceived the size of the risk reduction to be greater than did participants in the stacked bar graph condition, $t(399) = 4.11$, $p < .001$. Similarly, the perceived significance of the risk reduction was greater in the bar graph condition than in both the numbers condition, $t(399) = 3.48$, $p < .001$, and the stacked bar graph condition, $t(399) = 6.74$, $p < .001$. Participants in the numbers condition perceived the significance of the risk reduction to be greater than did participants in the stacked bar graph condition, $t(399) = 3.37$, $p = .001$.

Table 4 presents the means of participants' ratings of the considerations they used when deciding how much they were willing to pay for Improved Toothpaste. Participants in both the bar graph condition and the numbers condition were more influenced by the comparison between the number of people who developed gum disease (the foreground information: 30 versus 15) than were participants in the stacked bar graph condition, $t(387) = 3.16$, $p = .002$, and $t(387) = 3.15$, $p = .002$, respectively. The bar graph and numbers conditions did not differ with regard to this question, $t(387) = .078$, $p = .94$. Participants in the numbers condition based their

Table 4
Mean impact ratings for decision considerations by condition in Experiment 2

Dependent measure	Condition		
	Numbers	Bar graphs	Stacked bar graphs
Comparison of foreground information (30 vs. 15)	5.13 (1.64)	5.15 (1.75)	4.49 (1.64)
Comparison of foreground:background (30/5000 vs. 15/5000)	4.73 (1.71)	4.27 (1.86)	4.28 (1.90)

Note. Numbers in parentheses are standard deviations. Sample sizes varied per condition, but ranged from $n = 124$–135. Impact rating scores ranged from 1 to 7, where 1 = had no impact and 7 = had a major impact.

decision more on the comparison of the chances of developing gum disease with each toothpaste (foreground:background ratio) than did either participants in the bar graph condition $t(387) = 2.02$, $p = .04$, or the stacked bar graph condition, $t(387) = 2.01$, $p = .05$. Participants in the bar graph and stacked bar graph conditions did not differ in their responses to this item, $t(387) = .036$, $p = .97$.

3.3. Discussion

These results extend the findings of the first experiment by showing that an additional format that displays both the foreground and background information graphically eliminates the previously observed graphical effect as well. Unlike that study, however, although the trend was in the same direction, there was no reversal of the graphical effect regarding the amount participants were willing to pay for the safer product. Nevertheless, participants in the numbers condition did perceive a greater risk reduction than did participants in the stacked bar graphs condition. Given our hypothesis for why the previously observed graphical effect occurred, this result should perhaps not be surprising. If our proposed mechanism is correct, the effect of the graphical format on risk avoidance is mediated by the extent of the perceived risk reduction. Thus, it is reasonable that the influence of the display format would be greater on questions addressing the perceived risk reduction than on the behavioral measure of risk avoidance.

Additionally, we attempted to examine more directly participants' reasoning processes in this study as well. The foreground:background salience hypothesis suggests that participants should be more apt to state that they made their decisions by comparing only the foreground information when in the bar graphs condition than in either the numbers or stacked bar graphs condition. Similarly, participants in the bar graphs condition should be less apt to state that they made their decisions by comparing the foreground:background ratio than would participants in either the numbers or stacked bar graphs conditions. The results partially supported these predictions, in that participants presented with bar graphs said that their decisions were based on foreground comparisons more than did participants presented with stacked bar graphs, and less on foreground:background comparisons than did participants presented with numerical information. The other two relevant comparisons were not significant, however, suggesting that further work should be conducted with this procedure before strong conclusions can be made on the basis of it.

4. General discussion

The goal of the present research was to determine why the graphical techniques used by Stone et al. (1997) were effective at modifying risk-relevant behavior. Our experiments suggest the presence of a foreground:background salience effect, whereby the key factor is whether attention is called to the number of people at risk of harm (background information), or whether the focus is on the number of people harmed (foreground information). In support of this conclusion, two graphical formats not used by Stone et al. (1997) were developed that highlighted background information as well as foreground information. Both of these formats—a pie charts display in Experiment 1 and stacked bar graphs in Experiment 2—failed to increase risk-avoidant behavior to a greater extent than a numerical display. Additionally, on all of the relevant dependent measures (amount willing to pay, as well as the risk reduction measures), the order of the effects was the same: the display graphically highlighting only the foreground information produced the largest effects, the display graphically highlighting both the foreground and the background information produced the weakest effects, and the numbers display produced intermediate effects. The majority of these effects were statistically significant at the .05 level, with the only exceptions being between asterisks and numbers in Experiment 1 for willingness to pay and size of the risk reduction, and between numbers and the stacked bar graphs in Experiment 2 for willingness to pay. Even in these situations, however, the effects were in the same direction, and typically approached significance.

Thus, we are suggesting that the graphical effect identified by Stone et al. (1997) occurs because participants focus more on background information when presented with a numerical format displaying both foreground and background information numerically than when presented with a format displaying the foreground information graphically. This differential focus, then, produces a framing effect (see, e.g., Levin, Schneider, & Gaeth, 1998; Tversky & Kahneman, 1981), whereby different ways of presenting mathematically equivalent information produce different decisions. In particular, participants view the risk reduction as being larger when the foreground information only is displayed graphically than when both types of information are displayed either numerically or graphically. In other words, the attribute of risk reduction is seen as greater in the former condition than in the latter ones.[2] We have proposed that this differential focus on foreground versus background information occurs due to the level of saliency produced by the graphical display; nonetheless, it should be emphasized that we have never directly tested whether it is saliency that is responsible for the foreground:background effect. Further research is therefore needed to determine whether saliency or some other aspect of the display, such as evaluability (cf. Hsee, 1996), is responsible. Regardless of the exact form of the mechanism, however, it seems clear that displays that highlight foreground, but not background, information graphically will lead to the largest perceived risk reduction.

For the remainder of the paper, we will proceed as follows. First, we examine the question of whether purely numerical displays call attention equally to foreground and background information. Second, we discuss whether or not highlighting foreground information at the expense of background information should always produce greater risk-avoidant behavior. And finally, we end with a discussion of the practical implications of our findings.

4.1. Do numerical displays highlight the foreground and background equally?

As discussed, we suggest that participants in the Stone et al. (1997) experiments were more risk-avoidant when presented with graphical displays than with a purely numerical display because participants focused less on the background information when presented with graphical rather than numerical displays. Nonetheless, this finding does not necessarily imply that participants pay equal attention to foreground and background information when presented entirely with numerical information. Indeed, although Experiment 2 participants in the numbers condition said their decisions were influenced more by comparing the foreground:background ratio (30/5000–15/5000) than did participants in the bar graphs condition, they also stated that they were highly influenced by just comparing the foreground information (30–15). Thus, it appears plausible that participants presented with numbers were also more heavily influenced by the foreground information than by the background information, but just not to the same extent that participants presented with asterisks or bar graphs were.

---

[2] Although our results do indicate a type of attribute framing effect, it should be emphasized that this term is being used in a manner slightly different from that of Levin et al. (1998), who examined only valence framing effects in their review.

This possibility—that even in the numbers condition people attend more to foreground than to background information—is supported by a recent study by Yamagishi (1997). Testing a claim first made by Halpern et al. (1989), Yamagishi showed that a variety of events were rated as riskier when both foreground and background numbers were high than when both were relatively low. This finding suggests that, in the numbers condition, the foreground does affect decisions to a greater extent than does the background information. Nonetheless, it is clear that any additional salience associated with the foreground as compared to the background information in the numbers condition is not nearly as large as is the case with formats that display only the foreground information graphically. Indeed, in an earlier study similar to the one run by Yamagishi (1997), Halpern et al. found no effect of manipulating the magnitude of both foreground and background information.

If our reasoning is correct, then, the results of these studies may be highly dependent on precisely how the information is displayed to participants. Specifically, using formats that display only foreground information graphically in a design similar to that used by Yamagishi (1997) or Halpern et al. (1989) should produce stronger effects than found by either of those researchers. However, using formats that display both the foreground and the background graphically should eliminate that effect entirely.

4.2. Should highlighting foreground information always increase risk avoidance?

To the extent that our foreground:background hypothesis is correct, it should be clear that the graphical effect demonstrated by Stone et al., as well as the difference between the graphical conditions in our two experiments, is dependent on the incidence rates being small. If, for example, background information indicates that only a small number of people are at risk of harm, then calling attention away from this information should not have any effect on the perceived risk reduction. This line of reasoning is supported by a study by Stone and Rush (1997). In their research, Stone and Rush presented participants with a scenario stating that a company was considering marketing a running shoe that would reduce a certain risk. The risk was either a low-probability one (stress fracture to the heel: 30 per 4000 runners injured) or a high-probability one (knee pain: 30 per 65 runners affected), and participants were asked how much they would be willing to spend to reduce the risk in half. The graphical effect documented by Stone et al. (1997) was replicated in this experiment for the low-probability risk; however, there was no effect for the high-probability risk. The present research suggests that the failure to obtain the graphical effect with high-probability risks occurred because calling attention away from the background information no longer made the risk reduction seem larger. Instead, it appears that participants respond similarly to risk reductions of 30 to 15 as they do to risk reductions from 30 out of 65 to 15 out of 65.

4.3. Practical implications

The present ideas and findings have implications for risk communications quite generally, we believe, including in the domain of public health policy decision making.[3] Epidemiologists represent and communicate risks in a variety of ways (cf. Gordis, 1996; Palinkas & Hoiberg, 1982). The most basic representation is the *incidence*, or *incidence rate*, which is simply the proportion of times that the people at risk of being harmed by some hazard are indeed harmed. Thus, in our studies, the

[3] We thank an anonymous reviewer for encouraging us to examine this issue.

incidence of gum disease with Standard Toothpaste was 30/5000 = .006, and the incidence with Improved Toothpaste was smaller, 15/5000 = .003.

In public health discussions, the concern is frequently with comparisons of risks under alternative conditions, for example, of developing cancer when smoking cigarettes versus when not smoking. Hence, one of the most common risk representations in such discussions is the *relative risk (RR)*, which is the ratio of the incidences for the conditions at issue:

$$\text{RR} = (\text{Condition 1 Incidence})/(\text{Condition 2 Incidence}) \quad (1)$$

In the present experiments, the relative risk would be represented as

$$\text{RR} = (\text{Standard Incidence})/(\text{Improved Incidence}) = (30/5000)/(15/5000) = 2.0,$$

an indication that Standard Toothpaste is "twice as risky" as Improved Toothpaste. Another popular representation in public health, the *attributable risk (AR)*, focuses on the reduction in risk that could be achieved if there were a change from one condition to another or, equivalently, the "extra" risk that is attributable to the more hazardous condition:

$$\text{AR} = (\text{Condition 1 Incidence} - \text{Condition 2 Incidence})/(\text{Condition 1 Incidence}), \quad (2)$$

where Condition 1 is riskier than Condition 2. In our studies, we would have

$$\text{AR} = (\text{Standard Incidence} - \text{Improved Incidence})/(\text{Standard Incidence}) = (30/5000 - 15/5000)/(30/5000) = .5$$

implying that "half the risk experienced by Standard Toothpaste users is attributable to their use of that inferior toothpaste per se." In a public health discussion, the numerator of the attributable risk, Condition 1 Incidence - Condition 2 Incidence ("three one-thousandths" in our toothpaste scenarios), might be described along the lines of "the incidence of harm attributable to Condition 1" (see Gordis, 1996, p. 156, for more information on the interpretation of AR).

Note that, in both the RR and AR illustrations, the background frequency (5000) is cancelled out in the calculations. Thus, both the relative risk and the attributable risk effectively ignore background frequencies. Assuming the validity of the foreground:background hypothesis, then, presenting relative risk or attributable risk statistics should have similar effects to our formats that displayed only the foreground information graphically. Therefore, if the aim of public health communicators is to spur actions (say, policy decisions) that reduce risks in situations where absolute incidence rates are low, communicating in terms of relative and attributable risks, as is often done, should be an effective strategy. Conversely, describing the risks in terms of incidences of attributable harm takes background as well as foreground frequencies into account. Accordingly, similar to our formats that displayed both the foreground and background frequencies in the same form (either numerically or graphically), this approach would be less effective in inducing risk-avoidant behavior.

What is the normatively or even morally "right" thing for public health communicators (or anyone relying on these types of communications, such as product advertisers) to do? If one thinks of normativeness in terms of mathematical correctness, then the question is moot since all of these representations (incidences, relative risks, attributable risks, attributable incidences, displayed in numerical or graphical form) are formally coherent. But suppose that one's moral imperative is to induce people to take actions that protect them from harm while still being truthful. Then, using formats that call attention away from the background information, such as by displaying only foreground information graphically or using relative or attributable risk statistics, would be the appropriate solution.

Nonetheless, it is worth emphasizing that our work suggests that these techniques that would be most effective for increasing risk-avoidant behavior may actually do so by leading to a less complete understanding of the risks involved. That is, since the effectiveness of graphical formats that display only foreground information graphically (and other risk statistics like RR and AR) results from calling attention away from one of the two pieces of (relevant) information, it could be argued that using these types of graphical formats would result in a decreased understanding of the relevant risks. Instead, all the present work has shown is that calling attention away from the number of people at risk of harm is an effective technique for decreasing risk-taking behavior. A full discussion of whether or not modifying risk-relevant behavior without any commensurate increase in understanding of the risk levels involved is beyond the scope of the present paper (see Keeney & von Winterfeldt, 1986, for a useful discussion of this issue). However, two observations warrant mentioning. First, to the extent that it is generally accepted that excessive risk taking occurs, e.g., in terms of drinking while driving, then public policy initiatives that focus primarily on increasing risk avoidance in these situations may well be wise. Second, there is no reason to believe that being able to state relevant incidence rates indicates any true understanding of the risk levels involved.

This latter point is particularly important in the context of low probabilities. If, as suggested in prospect theory (Kahneman & Tversky, 1979) and other places, people often edit low probabilities to zero, there may be a common tendency to dismiss low-probability events out of hand. This could in turn produce the type of framing effect demonstrated by Slovic et al. (1978), in which participants were less apt to demonstrate risk-avoidant behavior when presented with risk information in terms of each driving trip than when the information was aggregated over a lifetime of driving. Thus, it appears plausible that by de-emphasizing background information, certain formats for conveying risk information serve to reduce people's tendency to dismiss low-probability events as never occurring.

Regardless, we acknowledge that the preceding arguments are speculative, and they by no means suggest that de-emphasizing background information is always legitimate for public health or other risk communications. As many researchers have pointed out, increasing risk avoidance is only one of many legitimate goals of risk-communication efforts. Thus, even when attempting to decrease risk taking, it is necessary to carefully consider how to meet that goal while at the same time focusing on other concerns, such as not misleading the recipients of the communication messages.

Acknowledgments

Partial support for this work was provided by the University of Michigan Business School. Portions of this research were previously presented at the November 1998 and November 2001 annual meetings of the Judgment and Decision Making Society. We are grateful for the assistance of Hannah Faye Chua and Laith Alattar in carrying out the work described here, and to Andrew Parker and Isaac Lipkus for helpful discussions regarding the ideas incorporated in this paper.

References

Baron, J. (1997). Confusion of relative and absolute risk in valuation. *Journal of Risk and Uncertainty, 14*, 301–309.

Brown, L. J., Oliver, R. C., & Loe, H. (1990). Evaluating periodontal status of US employed adults. *Journal of the American Dental Association, 121*, 226–232.

Camerer, C. F., & Kunreuther, H. (1989). Decision processes for low probability events: Policy implications. *Journal of Policy Analysis and Management, 8*, 565–592.

Covello, V. T., Sandman, P. M., & Slovic, P. (1988). *Risk communication, risk statistics, and risk comparisons: A manual for plant managers*. Washington, DC: Chemical Manufacturers Association.

Covello, V. T., von Winterfeldt, D., & Slovic, P. (1986). Risk communication: A review of the literature. *Risk Abstracts, 3*, 171–182.

Fischhoff, B. (1995). Risk perception and communication unplugged: Twenty years of process. *Risk Analysis, 15*, 137–145.

Fisher, A. (1991). Risk communication challenges. *Risk Analysis, 11*, 173–179.

Fisher, A., McClelland, G. H., & Schulze, W. D. (1989). Communicating risk under Title III of SARA: Strategies for explaining very small risks in a community context. *Journal of Air Pollution Control Association, 39*, 271–276.

Gordis, L. (1996). *Epidemiology*. St. Louis: W.B. Saunders.

Halpern, D. F., Blackman, S., & Salzman, B. (1989). Using statistical risk information to assess oral contraceptive safety. *Applied Cognitive Psychology, 3*, 251–260.

Howell, D. C. (1992). *Statistical methods for psychology*. Boston: PWS-Kent.

Hsee, C. K. (1996). The evaluability hypothesis: an explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67*, 247–257.

Jarvenpaa, S. L. (1990). Graphic displays in decision making—the visual salience effect. *Journal of Behavioral Decision Making, 3*, 247–262.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Keeney, R. L., & von Winterfeldt, D. (1986). Improving risk communication. *Risk Analysis, 6*, 417–424.

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

Lipkus, I. M., & Hollands, J. (2000). *The visual communication of risk. Monographs of the National Cancer Institute*. Bethesda, MD: National Institutes of Health.

Magat, W. A., Viscusi, W. K., & Huber, J. (1987). Risk-dollar tradeoffs, risk perceptions, and consumer behavior. In W. K. Viscusi, & W. A. Magat (Eds.), *Learning about risk* (pp. 83–97). Cambridge, MA: Harvard University Press.

Palinkas, L. A., & Hoiberg, A. (1982). An epidemiology primer: Bridging the gap between epidemiology and psychology. *Health Psychology, 1*, 269–287.

Reyna, V. F., & Brainerd, C. J. (1991). Fuzzy-trace theory and framing effects in choice: Gist extraction, truncation, and conversion. *Journal of Behavioral Decision Making, 4*, 249–262.

Rohrmann, B. (1992). The evaluation of risk communication effectiveness. *Acta Psychologica, 81*, 169–192.

Roth, E., Morgan, M. G., Fischhoff, B., Lave, L., & Bostrom, A. (1990). What do we know about making risk comparisons? *Risk Analysis, 10*, 375–387.

Sanfrey, A., & Hastie, R. (1998). Does evidence presentation format affect judgment? An experimental evaluation of displays of data for judgments. *Psychological Science, 9*, 99–103.

Siegrist, M. (1997). Communicating low risk magnitudes: incidence rates expressed as frequency versus rates expressed as probability. *Risk Analysis, 17*, 507–510.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1978). Accident probabilities and seat belt usage: A psychological perspective. *Accident Analysis and Prevention, 10*, 281–285.

Stone, E. R., Rush, C. J. (1997, November). Risk communication: *The effectiveness of graphical modes depends on the risk magnitude*. Poster presented at the annual meeting of the Society for Judgment and Decision Making, Philadelphia, PA.

Stone, E. R., Yates, J. F., & Parker, A. M. (1994). Risk communication: Absolute versus relative expressions of low-probability risks. *Organizational Behavior and Human Decision Processes, 60*, 387–408.

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.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453–458.

Viscusi, W. K., & Magat, W. A. (1987). *Learning about risk*. Cambridge, MA: Harvard University Press.

Weinstein, N. D., Kolb, K., & Goldstein, B. D. (1996). Using time intervals between expected events to communicate risk magnitudes. *Risk Analysis, 16*, 305–308.

Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: Implications for risk communication. *Applied Cognitive Psychology, 11*, 495–506.

Received 15 September 1999