Journal of Behavioral Decision Making
J. Behav. Dec. Making, **15**: 19–35 (2002)
DOI: 10.1002/bdm.401

Judging the Accuracy of a Likelihood Judgment: The Case of Smoking Risk

PAUL D. WINDSCHITL*
University of Iowa, USA

ABSTRACT

A standard method for assessing whether people have appropriate internal representations of an event's likelihood is to check whether their subjective probability or frequency estimates for the event correspond with the assumed objective value for that event. When a person's estimate for the event exceeds its assumed objective probability or frequency, the person's expectancy for the event is concluded to be greater than warranted. This paper describes three lines of reasoning as to why conclusions of this sort can be problematic. Recently published findings as well as data from two new experiments are described to support this main thesis. The case of smoking risk is used to illustrate the more general problem, and issues that must be considered to avoid or contend with the problem are discussed. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS subjective probability; frequency estimates; smoking risk

In a national survey of more than 3000 persons, respondents were asked: 'Among 100 cigarette smokers, how many do you think will get lung cancer because they smoke?' The mean response for this question was 43 (Viscusi, 1990). This mean is considerably higher than the true lung cancer risk, which Viscusi (1990) estimated to be between 5 and 10 per 100. From a finding like this, one could conclude that the survey respondents were oversensitive to the possibility of getting lung cancer from smoking. Indeed, data of this type have been used to support a key argument used by lawyers for the cigarette industry in civil litigation. They argue that smokers know the risks associated with smoking and are making a rational decision to smoke; therefore, smokers have no justification for suing if they become ill.

The fact that Viscusi (1990, 1991, 1992) interpreted his data as suggesting that people are oversensitive to lung cancer risk may not seem to be problematic. It is common for psychologists and other behavioral scientists to conclude that, if a person's probability or frequency estimate for a negative event exceeds the objective value, then the person's expectation that the event will occur is greater than warranted. The same is true for underestimations. For example, if a homeowner's numeric estimate for a significant radon exposure in his or her house is lower than what relevant base rates would suggest, it seems natural to conclude that the homeowner's expectation for the occurrence of radon exposure is less than it should be. The main thesis of the present paper, however, is that this type of conclusion is often problematic. More specifically, this paper

*Correspondence to: Paul D. Windschitl, Department of Psychology, University of Iowa, Iowa City, IA 52242, USA.
E-mail: paul-windschitl@uiowa.edu

concerns conclusions that are drawn from comparisons between subjective probability or frequency estimates for a specific event and an appropriate objective standard for that event. I argue that although such a comparison is obviously relevant for judging the accuracy of a person's numeric estimate, the comparison should not necessarily be used as the basis for concluding whether the person is appropriately sensitive to the potential for the target event.

In this paper, I describe three interrelated reasons—each with accompanying lines of research—as support for the proposal that comparisons between subjective and objective estimates for an event's probability or frequency can be problematic. The first of the three reasons concerns evidence indicating that respondents often misuse or misinterpret numeric probability measures. Many of these misuses have the potential to systematically shift the mean estimate for a group of respondents, thus clouding the interpretation of the data. Turning to the second reason, I extensively discuss several recent lines of research that provide evidence for a novel distinction between people's reports of numeric probability and their more intuitive perceptions of an event's likelihood. An important potential consequence of this distinction for the smoking-risk issue is that even if a smoker believes his or her numeric probability of dying from lung cancer is 0.45, he or she can still have a very high sense of optimism—at an intuitive, nonanalytic level—about not dying from lung cancer. With respect to the third reason, I discuss the idea that subjective probability and frequency responses are often *ad hoc* constructions (not pre-formed beliefs) that are based on information that is, at the time the response is made, highly salient and accessible. Therefore, an individual estimate or a mean estimate from a group can be highly dependent on question framing and context factors. Whereas existing research is cited and described to support the first two reasons outlined here, two new experiments were conducted to test hypotheses relevant to the third reason—specifically within the domain of smoking risk. Despite my cautionary emphases about drawing conclusions from comparisons between subjective and objective estimates about a single event, I do not intend to suggest that biases/accuracies in respondents' probability and frequency estimates cannot be assessed; the concluding section of this paper discusses conditions in which subjective estimates of probability or frequency can more justifiably be used to make judgments of whether people have an appropriate expectation about the possibility of an event.

Before discussing the three reasons underlying the paper's main thesis, I should note that I do not attempt to summarize in this paper the many facets of the debate regarding Viscusi's (1992) arguments or smoking risks in general. The reader is referred to recent papers that address and cite relevant literature on issues such as whether smokers underestimate the ease with which they would be able to quit, whether smokers can accurately predict their smoking rates at a future point in time, whether smokers minimize their own personal vulnerability relative to other smokers, whether smokers consider the cumulative risks of smoking, and what is the best objective estimate of lung cancer risk for smokers (see Hanson and Kysar, 1999; Slovic, 2000a,b; Viscusi, 1992, 2000; Weinstein, 1998). The present paper focuses on comparing subjective and objective estimates for a single event and how concerns about this practice have implications for the interpretation of Viscusi's data.

REASON 1: RESPONDENTS' MISUSE AND MISINTERPRETATION OF THE SUBJECTIVE PROBABILITY MEASURE

In a recent paper on perceptions of smoking risks, Weinstein (1998) argued that people have difficulty using the odds and percentages that are often employed to assess perceived certainty of various health outcomes. I wish to echo and extend this argument as an important reason why comparisons between subjective and objective probability estimates for a single event—in any domain—can be problematic.

While many people learn something about probability theory in their educational experience, it is difficult to know the extent to which they internalize that information for future use outside of mathematics and statistics classes. In everyday contexts, most people have little or no experience with generating numeric probability estimates and receiving systematic feedback about accuracy. Behavioral scientists, on the other hand, typically have extensive training in statistics. On nearly a daily basis, they are reminded of and use probability values as a way of expressing likelihood and confidence in research findings. Given the vast difference in training and experiences between behavioral scientists and most individuals, it is risky when researchers simply assume that their research participants have or can quickly develop an adequate understanding of the properties of a probability scale in order to effectively communicate their internal perceptions of events' likelihoods. While researchers have little trouble recognizing, for example, the ratio-scale properties of probabilities or the additivity constraints placed on a set of mutually exclusive and exhaustive events, research participants may have different understandings.

There are two types of use/interpretation problems that can be distinguished. The first problem, one which has close ties to the psychophysics literature (Poulton, 1989), concerns systematic biases in the way in which research participants use the metric specified by the researcher. If the mean subjective estimate for a specified dimension of a stimulus exceeds its objective value, we can appropriately say that participants overestimated the value. However, without examining the mean subjective estimates given for other objective values, it is difficult if not impossible to know whether the overestimate is due to a systematic bias in the way in which the particular metric was used, or whether the overestimation occurred because of a flawed internal representation of the specific value or magnitude of interest to the researcher.

As an analogy, imagine that we ask a group of college basketball players from the USA to estimate the distance between a section of the three-point line and the point on the floor directly beneath the hoop—in centimeters. Their responses might greatly overestimate (or underestimate) the actual distance. If so, should we assume that these players don't have a good appreciation of the distance between the three-point line and the basket? No; given a blank floor these players may well be able draw the three point line within a few centimeters of its actual location. In this example, it is easy to see that the overestimation in centimeters is more likely due to an imperfect use of the metric than a flawed mental representation of the actual distance. Switching the metric from centimeters to feet would likely help the accuracy of the estimates, since basketball players in the USA have much more experience measuring, estimating, and getting feedback about lengths in feet than centimeters (and biases tend to be reduced as the familiarity of a response scale increases; Poulton, 1989). However, even with a metric as familiar as the foot, a researcher would not assume that an overestimate necessarily was evidence of a flawed mental representation of distance from the three-point line to the hoop. Separating response biases from flawed mental representations would require that a researcher solicits judgments for stimuli that differ in objective values.

Similarly, risk perception researchers should recognize that an overestimation of an event's numeric probability does not necessarily constitute evidence that the respondent is overly expectant about the possible occurrence of that event. If probability judgments about other unrelated events were also assessed, the meaning of the overestimation for the target event could change dramatically depending on whether probabilities for other events were also overestimated (suggesting a broad bias) or if they, unlike the target event, were estimated relatively accurately.

The second type of problem, which can be distinguished from the problem of systematic bias in the use of the metric, concerns the potential for respondents to *misinterpret* key aspects or constraints of the subjective probability measure. Although I do not wish to argue that the bulk of research participants have fatally flawed understandings of the 0–100% scale, a variety of research clearly demonstrates that a notable portion of participants use the 0–1.0 or 0–100% scale in ways not intended by researchers or in ways that indicate the participants do not understand the constraints of the scale (e.g. Black, 1995; Borland, 1997; Fischhoff and Bruine de Bruin, 1999; Teigen, 1974; Windschitl, 2000). For example, Fischhoff and Bruine de Bruin (1999) showed that, in response distributions in many data sets, there is a 'blip' at the 50% mark. In other words, the frequency of the 50% responses is peculiarly high. Fischhoff and Bruine de Bruin argue that for many participants, the 50% response is used as an indication that they have no idea as to the answer. Additional research demonstrated that use of the 50% response is greatest among less numerate respondents, children, and less educated adults, and also when the event in question evokes feelings of little perceived control (Bruine de Bruin *et al.*, in press). Another example of problems in people's use of probability scales comes from studies investigating the additivity of subjective probability estimates, which have revealed that participants' responses for a set of mutually exclusive and exhaustive events often greatly exceed 100% (e.g. Robinson and Hastie, 1985; Teigen, 1983; Tversky and Koehler, 1994; Wright and Whalley, 1983). This error is especially noteworthy in studies in which participants provided probability estimates for each event in the set in a successive fashion, making it possible for the participants to monitor the additivity of their responses (e.g. Teigen 1974; Windschitl, 2000). In such studies, the error suggests that participants have either a lack of awareness or a blatant disregard for the constraints of a numeric probability scale. Do the above-mentioned misunderstandings or misuses of a subjective probability scale mean that it cannot be used to meaningfully assess perceptions of likelihood? No, but they do indicate that a probability estimate for any one event (or the mean of the estimates from a group) may be inflated or deflated by factors that are not directly relevant to the construct that the researcher is attempting to assess.

Finally, in a paper on smoking risk, Borland (1997) discussed two distinct ways in which participants might interpret a probability measure: (1) as a measure of concern, or (2) as a standard measure on which a response reflects an estimate of the 'objective' probability or relative frequency of the event (see also Eiser, 1994). Borland (1997) pointed out that, when a probability measure is treated as a measure of concern, it would make sense that a response (i.e. a concern response) regarding a disease that is the most likely disease for a smoker would be located on the upper half of the scale, even if that disease's actual probability was less than 50%. Hence, the concern interpretation makes it questionable for a researcher to make judgments of whether participants' estimates of probability matched true probabilities; after all, the participants using that interpretation didn't provide probability estimates but rather concern ratings. Borland's point is a critical one, especially because a researcher's interpretation of a mean probability estimate for an event can be seriously affected even if only a portion of the participants adopt a concern interpretation of the probability measure. Future research should be done to determine whether, when, and how often people assume the concern interpretation, as well as why and who is most likely to assume this interpretation.

REASON 2: BELIEFS IN OBJECTIVE PROBABILITY AND INTUITIVE PERCEPTIONS OF CERTAINTY

The preceding section described problems with how respondents interpret and use probability measures. If a researcher can rule out these problems and assume that respondents are adequately familiar with probability measures, are comparisons between subjective and objective probability estimates for a specific event warranted? This section describes a different type of problem that applies to these comparisons, even when all the problems associated with the interpretation and use of the subjective probability scale can be ruled out.

The second reason for suggesting caution about comparisons between subjective probabilities and the relevant objective probability for an event is that numeric subjective probabilities might not be tapping into the certainty representations and processes that mediate the decisions or behaviors that are ultimately of interest to the researcher. It has typically been assumed, with some notable exceptions (e.g. Gigerenzer, 1994; Howell and Burnett, 1978; Howell and Kerkar, 1982; Teigen and Brun, 2000; Zimmer, 1983), that different measures of uncertainty are largely interchangeable and that subjective probability measures serve as a precise, relatively reliable, and convenient way of assessing the internal construct of uncertainty. Recent research, however, has provided a variety of evidence, described in detail below, that a person's belief about the numeric probability of an event is not the only critical component to uncertainty. Although this belief may play a key role in mediating decisions and behavior, and although a numeric probability measure would be the best way of assessing that belief, there is a more intuitive and non-analytic component to uncertainty that is not necessarily well represented in a numeric subjective probability response but can be an important mediator of decisions and behavior.

The evidence for this position comes from a variety of studies and will be described here in two categories. One type of evidence comes from recent studies investigating alternative ways of assessing people's certainty about events for which a rule-based and analytic assessment of relevant information might produce a different judgment about certainty than would a more intuitive and non-analytic assessment. In a paper that examined numeric likelihood measures and verbal likelihood measures (which included response options like 'quite likely', 'fairly unlikely'), Windschitl and Wells (1996) argued that numeric measures prompt people to think about a target event in a more deliberative and rule-based fashion than they typically would. Verbal measures, alternatively, are less likely to prompt this type of thinking and therefore are more sensitive to the intuitive, associative, and non-analytic thoughts about certainty that mediate decisions and behaviors relevant to the event (for information on the general distinction between rule-based and associative processes, see Epstein, 1990, 1994; Sloman, 1996; Smith and DeCoster, 1999, 2000). This line of reasoning was supported in two experiments in which participants read information about target events, provided either numeric or verbal likelihood estimates for those events, and also indicated preferences or behavioral intentions that were relevant to the uncertain outcomes of those events. As expected, numeric certainty estimates, although sensitive to the rule-based information presented about the events (e.g. base-rate information), did a poorer job of predicting participants' preferences and behaviors than did verbal certainty estimates.[1]

Similar evidence comes from studies on the alternative-outcomes effect (Windschitl and Wells, 1998; Windschitl and Young, in press; see also Gonzalez and Frenck-Mestre, 1993). In one such study, participants provided either numeric or verbal certainty estimates for whether a chocolate-chip-loving girl, who was blindly drawing a cookie from a jar, would get her desired cookie (Windschitl and Wells, 1998). Some participants read that the jar contained 2 chocolate chip cookies and 7 oatmeal cookies, whereas others read that it contained 2 chocolate chip, 1 oatmeal, 1 raisin, 1 butterscotch, 1 rum, 1 peanut butter, 1 pecan, and 1 sugar cookie. Numeric probability responses showed no significant sensitivity to this alternative-outcomes manipulation, but verbal certainty responses revealed that participants felt significantly more optimism in the latter version of the scenario. In related work on the equiprobability effect, Teigen (in press) demonstrated that when a focal outcome is equal in probability to all alternative outcomes, people will often describe the focal outcome with verbal probability phrases that clearly connote optimism, even when there are several alternative outcomes and the focal outcome is acknowledged to have only a small numeric probability (e.g. 16.6% for a six-outcome situation). It is important to note that additional evidence from Windschitl and Wells (1998) and Teigen (in press) demonstrated that these manipulations to the distributions of evidence for alternative outcomes have robust effects not only on verbal measures of certainty, but on a variety of relevant decisions and behaviors as well. Hence, the variability in perceived certainty detected by the verbal measures and missed by the numeric measures was not spurious; instead it was an important mediator of subsequent responses.

A second related type of evidence for a distinction between beliefs in objective probabilities and more intuitive perceptions of certainty comes from research in which participants are made explicitly aware of the numeric probability of an event, but at the same time are provided with information that might affect their reactions, judgments, or intuitive perceptions about the probability value. A relevant line of research comes from Epstein and his colleagues on the ratio-bias phenomenon (see Denes-Raj and Epstein, 1994; Denes-Raj *et al.*, 1995; Kirkpatrick and Epstein, 1992; Pacini and Epstein, 1999). In one key study, participants were given a choice between two bowls from which they would attempt to draw a winning bean (Denes-Raj and

[1]For a discussion of when verbal measures of certainty are unlikely to be superior to numeric measures, see Windschitl and Wells, 1996. For discussions of related research comparing verbal and numeric likelihood expressions on a variety of dimensions, see, for example, Budescu and Wallsten, 1995, Ofir and Reddy, 1996, Wallsten *et al.*, 1993.

Epstein, 1994). The bowls differed in proportions of winning beans and absolute numbers of winning beans. On one trial, for example, one bowl offered a 10% chance of winning but a small absolute number of winners (e.g. 1 in 10) while the other bowl offered a smaller likelihood of winning but a larger absolute number of winners (e.g. 7 in 100). Frequently, participants would choose to select from the bowl that contained a greater absolute number of winners, even though they were fully aware that the other bowl offered a greater chance of winning. Hence, belief in objective probability was not the only type of uncertainty construct that was guiding the choice behavior; some form of intuitive or non-analytic processing was sensitive to the absolute numbers of winning beans in the two bowls. Findings that are conceptually related to the ratio-bias phenomenon but in more applied domains were described by Slovic *et al*. (2000) and Yamagishi (1997).

Other relevant research has investigated how irrelevant context information can influence reactions to precise numeric probability estimates from forecasters (Windschitl and Weber, 1999). One scenario used in this research described a woman who was told by her doctor that, because of a diagnosed blood condition, she would have a 30% chance of developing a disease related to malaria on an upcoming trip. Participants who read that the doctor's 30% estimate referred to a trip to India expressed greater certainty on a nonnumeric scale that she would get the disease than did participants who read that the 30% estimate referred to a trip to Hawaii. Windschitl and Weber argued that although participants in both groups believed the 30% forecast was an accurate estimate of probability, their more intuitive perceptions of certainty were influenced by the strength of the associative links between malaria-like diseases and India/Hawaii.

Relatedly, in a study on how social comparisons influence perceptions of vulnerability, Klein (1997) asked participants to imagine they had just been tested for a genetic marker of a pancreatic disease. They were then given hypothetical information about their own chances for developing a pancreatic disease as well as the chances for the average person of the same age and sex. Participants' responses to the question of how disturbed they would be about this information were significantly affected by a manipulation of the relative risk information (about how they compared to the average person) but not a manipulation of the absolute risk information (whether their risk was 30% or 60%). In fact, participants who imagined that they had a 60% risk but thought that this risk was lower than average felt less disturbed than participants who imagined they had a higher-than-average 30% risk. In another study on perceptions of vulnerability, Windschitl *et al*. (in press) demonstrated that when people learn about a disease prevalence rate (or frequency rate) for specified target group (e.g. 12% prevalence among women), their intuitive perceptions of how vulnerable this target group is to the disease differ depending on the prevalence rate for the same disease in a comparison group (e.g. men).

Taken together, the studies described in this section provide compelling evidence that people's beliefs in objective probability are at least partially independent of their intuitive perceptions of certainty. Numeric subjective probability measures, which are particularly suited for assessing beliefs about objective probability, may not always adequately reflect how the person intuitively perceives the likelihood of the event. This conclusion would be most applicable to cases in which people have a mix of information—some amenable to rule-based, analytic thinking and some amenable to more intuitive, non-analytic thinking—on which to base their judgment. These conditions appear to be met in the case of the perceived certainty of smoking hazards. Information about the risks/benefits of smoking comes in many forms, ranging from abstract and pallid statistics to magazine images of healthy and attractive individuals smoking. Hence, if a smoker estimates his or her chances of dying from lung cancer to be 40%, we can appropriately conclude that his or her 40% estimate is an overestimate of the best objective standard, but it is not possible to know from this number whether the person's more intuitive and non-analytic perceptions of vulnerability to lung cancer are "in line" with what they should be. The research described above suggests that there may be a variety of non-analytic considerations that might cause the person to feel relatively optimistic or pessimistic about this response (or in spite of this response). If the smoker's behavior is largely being mediated by intuitive and non-analytic considerations, the 40% overestimate has very limited meaning for the researcher interested in smoking decisions and behaviors.

REASON 3: PROBABILITY RESPONSES AS *AD HOC* BELIEFS

An underrecognized property of subjective probability judgments is that they are typically *ad hoc* constructions. That is, they are not beliefs that are stored *per se* in memory, but rather are constructed because they were solicited by the researcher. A given response to the question 'How likely are you to contract lung cancer?' may be influenced in part by some stored statistics about the prevalence of lung cancer, but the response will also be a function of associative links that are triggered by and most active in a current or recent context (see Eiser, 1994). Hence, like many constructs of interest to behavioral scientists—such as self-reports of behavior (e.g. Schwarz *et al.*, 1985) and estimates of hazard severity (Eiser and Hoepfner, 1991)—reports of subjective probability are susceptible to variations in the way in which questions are asked, the context in which the questions are embedded, and the way in which the response options are depicted. For example, Slovic and Monahan (1995) recently illustrated how the spacing of subjective probability response options can have a robust impact on clinicians' judgments about the probability that patients hospitalized for a mental disorder would harm someone if released (see also Slovic *et al.*, 2000). Such a finding suggests that even when a researcher has ensured that his or her participants understand how the probability scale is to be used and the researcher has confidence that the measure is tapping into the intended construct, a third potential problem exists: the researcher's choice of question wording and context can systematically shift mean likelihood estimates. When two or more questions about the same target event can produce substantially different mean subjective likelihood estimates because of variations in wording and context, it is problematic to treat the mean estimates from any one question as a reflection of a group's stable, well-formed beliefs about the likelihood of the event.

To what extent are questions about the probabilities of smoking-related hazards susceptible to wording and context effects? Recall that the question that Viscusi (1990, 1991, 1992) used to draw conclusions about people's perceptions of smokers' risks of lung cancer was: 'Among 100 cigarette smokers, how many do you think will get lung cancer because they smoke?' The mean response was 43. Below, I report two experiments to examine the impact of wording and context variations on responses to a question similar to Viscusi's. If these factors have a substantial impact on mean estimates, it would be difficult to argue that people carry stable, well-formed beliefs about the numeric probability of smokers' dying from lung cancer.

Beyond the question of whether wording and context would influence responses was the question of whether the impact of these factors would be roughly equivalent for both smokers and nonsmokers. One could argue that smokers, when making an informed and rational decision about smoking (see Viscusi, 1992), have already done considerable thinking about the chances of dying from lung cancer because of smoking (and perhaps they frequently reconsider these chances). If so, smokers in particular may have well-formed beliefs about their own personal likelihood of dying from lung cancer. It was expected, however, that even for smokers, responses to subjective probability questions about personal vulnerability do not reflect stable beliefs about numeric probabilities, but are largely a function of ad hoc thinking and can be substantially influenced by wording and context manipulations.

EXPERIMENT 1

The first experiment examined the role of explicitly listing, and having participants respond to, alternatives to lung cancer. Although the wording of Viscusi's (1990) question may seem relatively unbiased, it included only one potential cause (smoking) and only one potential effect (lung cancer). Hence, while one consequence of smoking is explicitly mentioned in the question, it was left to the respondent to recall other possible effects and noneffects of smoking. Does this type of question boost the perceived likelihood of a smoker getting lung cancer relative to a question that requires participants to consider alternatives to lung cancer?

Research related to a prominent theory of probability judgment called support theory suggests that the explicitness of alternatives to lung cancer is likely to play a role in the judged probability of lung cancer (Rottenstreich and Tversky, 1997; Tversky and Koehler, 1994). This research has demonstrated that as alternatives to a focal event are unpacked (i.e. explicitly represented in the critical question), the perceived support or evidence for those nonfocal alternatives increases and the judged probability of the focal event decreases (e.g. Fischhoff *et al.*, 1978). When alternatives are left in a packed form (or not explicitly represented in the question) they are less likely to be recalled and will have less impact on the judged probability of the focal event, which by necessity is always explicitly represented in the question.

Slovic (2000b) recently tested whether the implicit/explicit representation of alternatives has a similar effect for responses to a lung-cancer question (see also Hanson and Kysar, 1999; Slovic, in press). The university students in his experiment were asked on the first page of a survey to indicate how many of 100 smokers (1 pack per day) 'are likely to die of lung cancer'. Then on the second survey page, the same students were asked to indicate the numbers of these 100 smokers who were likely to die from each of 15 causes of death, one of which was lung cancer. The vast majority of respondents lowered their estimates for lung cancer from page one to page two. While the mean estimate on page one was 56, the mean estimate for lung cancer on page two was substantially lower at 20. Furthermore, the correlation between participants' page one and page two estimates was only 0.21, indicating little stability in the estimates.

Experiment 1 of the present paper was similar to Slovic's but with two notable differences. First, whereas Slovic used a within-subjects design, which likely enabled participants to become aware of the manipulation being tested, this experiment used a completely between-subjects design. Second, instead of answering a frequency question about smokers dying from lung cancer, half the participants in the present experiment answered a probability question about their own chances of dying from lung cancer. This allowed for a test of whether smokers' estimates of their own probability of dying from lung cancer was affected by whether alternative causes of death were represented implicit or explicitly. In line with the reasoning described above, it was expected that responses from smokers and nonsmokers alike would show sensitivity to this manipulation.

Method

Participants and procedure

The participants were 577 students from introductory psychology courses at the University of Iowa. In large group testing sessions, they completed numerous personality questionnaires and a short survey that was constructed for this experiment.

Design and materials

The design for the experiment was a 2 $\times$ 2 between-subjects design. Each participant responded to a survey that was either a probability or frequency version and either an implicit- or explicit-representation version.

The implicit-representation version of the probability question asked 'What is the chance that you will die from lung cancer? Place a chance estimate between 0% and 100% in the space below.' The explicit-representation version of probability question (shown in the Appendix) asked, 'What is the chance that you will die from each of the following?. . .Because the list includes all possible causes of death, and because a person's death is attributed to only one cause, your responses should add up to 100%.' Participants reading this explicit-representation version gave probability estimates for nine causes and an 'all other possible causes' category.

The implicit version of the frequency question asked 'Among 100 people who smoke 2 packs per day, how many would die from lung cancer?' The explicit version of the frequency question (shown in the Appendix) asked 'Among 100 people who smoke 2 packs per day, how many would die from each of the following?

Because the list includes all possible causes of death, and because a person's death is attributed to only one cause, your responses should add up to 100.' Participants reading this explicit version gave frequency estimates for the same causes that were listed in the probability version. At the bottom of all surveys, respondents were asked to indicate the average number of cigarettes that they smoked per week.

There were actually two explicit-representation versions of the probability survey, which differed only in the order in which the causes of death were listed (which was the reverse from what is shown in the Appendix). The same was true for explicit-representation versions of the frequency surveys. This manipulation had no effect on the lung cancer estimates and the relevant cells were collapsed together for the analyses reported here. One consequence of collapsing the cells is that the cell sizes for the explicit-representation cells are twice the size of those for the implicit-representation cells. However, analyses treating this as a six cell design produce identical conclusions to those based on the analyses presented below.

Results and discussion

Exhibit 1 displays the means and other statistics for the responses of all 577 participants to the lung cancer question. Exhibit 2 displays the same statistics but only for the 175 participant who were smokers (i.e. those participants who indicated that they smoked an average of one or more cigarettes per week). As a starting point for discussing the results, consider the mean response of participants who answered a question that most resembled Viscusi's (1990) survey question—that is, a frequency question that did not explicitly represent any alternatives. The mean response of these participants was 48. Although this question asked about the

Exhibit 1. Statistics based on responses from all participants to the lung cancer question in Experiment 1

Question type Representation of alternatives	Mean	(SD)	Median	*N*
Probability question				
Implicit	15.2	(17.6)	10.0	100
Explicit	9.6	(11.2)	5.0	194
Frequency question				
Implicit	48.0	(28.3)	50.0	91
Explicit	30.4	(18.4)	25.0	192

Note: The probability question asked the respondent to indicate his/her chance of dying from lung cancer, whereas the frequency question asked the respondent to indicate how many of 100 smokers would die from lung cancer. The explicit-representation version, but not the implicit-representation version, listed alternatives to lung cancer.

Exhibit 2. Statistics based on smokers' responses to the lung cancer question in Experiment 1

Question type Representation of alternatives	Mean	(SD)	Median	*N*
Probability question				
Implicit	29.2	(22.5)	20.0	28
Explicit	17.3	(14.3)	10.0	57
Frequency question				
Implicit	47.3	(29.1)	50.0	25
Explicit	31.0	(16.3)	30.0	65

Note: These data are from smokers only (i.e. participants who indicated that they smoke one or more cigarettes per week). The probability question asked the respondent to indicate his/her chance of dying from lung cancer, whereas the frequency question asked the respondent to indicate how many of 100 smokers would die from lung cancer. The explicit-representation version, but not the implicit-representation version, listed alternatives to lung cancer.

frequency of death from lung cancer, whereas Viscusi's question asked about simply getting lung cancer, the mean responses from this survey (48) and Viscusi's (43) are, roughly speaking, in the same ballpark.

The initial question of interest in this experiment is whether responses on the frequency and probability questions about lung cancer were significantly influenced by whether alternative causes of death were explicitly represented. As one can infer from Exhibit 1, the answer is 'yes' for both the frequency and probability versions. Like the results of Slovic's study (2000b), the mean estimate for lung cancer frequency was substantially higher when no alternatives were explicitly represented ($M = 48$) than when a set of alternative causes of death were explicitly represented and considered by respondents, ($M = 30$), $t(281) = 6.3$, $p < 0.001$. The same type of effect was found for responses to the probability questions, which asked about personal likelihood of dying from lung cancer, $t(292) = 3.3$, $p = 0.001$.

Perhaps more interesting is the question of how smokers' estimates of their own chances of dying from cancer are influenced by the implicit/explicit manipulation. As mentioned above, one might propose that smokers have already given considerable thought to the likelihood of lung cancer. If, as Viscusi (1992) suggests, smokers are well aware of the risks of lung cancer and have actually considered the probability of lung cancer in a rational decision process, then their probability responses should be relatively impervious to question-representation manipulations. The results, however, indicate that smokers' estimates for their own chances of dying from lung cancer were substantially influenced by the implicit/explicit manipulation (means = 29.2 and 17.3, respectively), $t(83) = 3.0$, $p < 0.01$). Smokers' estimates on the frequency question showed the same type of effect (means = 47.3 and 31.0), $t(88) = 3.4$, $p = 0.001$. There were no indications from interaction terms that the effect of the implicit/explicit manipulation was stronger for nonsmokers than for smokers.

EXPERIMENT 2

Experiment 1 demonstrated how, for smokers and nonsmokers alike, responses to frequency and probability questions about lung cancer death can be substantially shifted by manipulating whether the alternatives to lung cancer are implicitly or explicitly represented. Experiment 2 tests whether responses to the lung cancer question can be affected to a similar degree by another type of manipulation—one that varies the context preceding the target question rather than aspects of the target question itself. More specifically, the experiment tested whether responses to a target question about half-a-pack-per-day smokers dying from lung cancer would be influenced by whether the preceding question asked about nonsmokers dying from lung cancer or 2-packs-per-day smokers dying from lung cancer.

Method

Participants and procedure

A separate sample of 572 University of Iowa students participated in Experiment 2. In large group testing sessions, they completed numerous personality questionnaires and a short survey that was constructed for this experiment.

Design and materials

The design for the experiment was a 2(probability/frequency) $\times$ 2(non/heavy smoker context) between-subjects design. The target question in the probability version of the survey was 'What is the chance that a person who smokes half a pack per day would die from lung cancer?' In the nonsmoker-context version, this question was immediately preceded by 'What is the chance that a person who doesn't smoke would die from lung cancer' In the heavy-smoker-context version, this question was immediately preceded by 'What is the chance that a person who smokes 2 packs per day would die from lung cancer?'

The target question in the frequency condition was 'Among 100 people who smoke half a pack per day, how many would die from lung cancer?' Analogous to the probability versions, the question preceding the target question solicited frequencies for either nonsmokers or 2-pack smokers. At the bottom of all surveys, respondents were asked to indicate the average number of cigarettes that they smoked per week.

Results and discussion

As one would expect, participants provided higher responses to the context question itself if that question was the heavy-smoker-context question than if it was the nonsmoker-context question, $F(1, 568) = 612.9$, $p < 0.001$.[2] The key question of interest, however, was whether the type of context question influenced responses on the subsequent target question about lung cancer probability/frequency rates for half-pack-per-day smokers. Exhibit 3 displays the means and other statistics for the responses of all 572 participants to the target question. As can be inferred from the table, the context manipulation had a substantial main effect on responses, $F(1, 568) = 44.4$, $p < 0.001$. The interaction was not significant, suggesting that the effect of the context manipulation was about the same regardless of whether participants were providing probabilities or frequencies. Manipulating the context question that preceded the target question caused the mean of the probability responses for the target question to shift by 13.7 units (i.e. percentage points), and it caused the mean of the frequency responses to shift by 13.7 units (i.e. deaths).

The main effect for response type was also significant, $F(1, 568) = 21.3$, $p < 0.001$. This type of main effect was not meaningful in Experiment 1 because question content covaried with response type (i.e. the frequency question asked about smokers' chances of dying from lung cancer whereas the probability question asked about the respondent's chances). However, in Experiment 2, question content was roughly equivalent in the frequency and probability versions. Hence, the response-type main effect can be interpreted as another illustration of the point that probability and frequency questions are often not equivalent methods of measuring uncertainty and can produce disparate results (see e.g. Griffin and Buehler, 1999; Harries and Harvey, 2000; Howell and Burnett, 1978; Howell and Kerkar, 1982; Gigerenzer, 1994; Gigerenzer and Hoffrage, 1995, Slovic and Monahan, 1995; Slovic *et al.*, 2000).

Exhibit 4 displays statistics for the responses of only the 207 participants who were smokers. The results for smokers were essentially the same as those for the entire sample. The context and response-type main effects were significant, and the interaction was not significant; $F(1, 203) = 16.1$, $p < 0.001$; $F(1, 203) =$

Exhibit 3. Statistics based on responses from all participants to the lung cancer question in Experiment 2

Question type Preceding context	Mean	(SD)	Median	*N*
Probability question				
Nonsmoker context	57.0	(24.3)	60.0	142
Heavy smoker context	43.3	(23.6)	42.0	145
Frequency question				
Nonsmoker context	47.5	(26.4)	50.0	144
Heavy smoker context	33.8	(24.1)	30.0	141

Note: The target question asked about the probability or frequency of half-pack-per-day smokers dying from lung cancer. In the nonsmoker-context condition, the preceding question asked about nonsmokers. In the heavy-smoker-context condition, the preceding question asked about 2-pack-per-day smokers.

[2]The mean responses to the heavy-smoker and nonsmoker probability questions were 66.8 and 23.2, respectively. The mean responses to frequency versions of those questions were 54.3 and 13.3, respectively. The response-type × context interaction was not significant ($F < 1$), but the response-type main effect was significant, $F(1, 568) = 43.0$, $p < 0.001$.

Exhibit 4. Statistics based on smokers' responses to the lung cancer question in Experiment 2

Question type Preceding context	Mean	(SD)	Median	*N*
Probability question				
Nonsmoker context	53.8	(24.2)	57.5	44
Heavy smoker context	40.5	(21.6)	40.0	53
Frequency question				
Nonsmoker context	44.4	(26.5)	47.5	60
Heavy smoker context	31.1	(21.5)	27.5	50

Note: These data are from smokers only (i.e. participants who indicated that they smoke one or more cigarettes per week). The target question asked about the probability or frequency of half-pack-per-day smokers dying from lung cancer. In the nonsmoker-context condition, the preceding question asked about nonsmokers. In the heavy-smoker-context condition, the preceding question asked about 2-pack-per-day smokers.

8.01, $p < 0.01$; $F(1, 203) = 0.0$, $p > 0.05$; respectively. In a 2×2 ANOVA that included smoking status as a dichotomous factor, the smoking status factor exhibited a significant main effect but no significant interaction effects. Hence, smokers exhibited roughly the same sensitivity to context manipulations as did nonsmokers.

CONCLUSIONS FROM EXPERIMENTS 1 AND 2

I argued that a third reason for caution when comparing subjective and objective estimates for an event's probability is that subjective estimates are often not a reflection of stable, well-formed beliefs but instead are ad hoc constructions that are highly dependent on current and recent context. Experiments 1 and 2 were conducted to test the extent to which this argument applies to the lung cancer question used by Viscusi (1990). The results showed that respondents' estimates about the probability or frequency of smokers' lung cancer deaths depended dramatically on whether the question soliciting the estimates included alternatives to lung cancer and on the nature of the question immediately preceding the target question. The effects were just as strong for smokers as they were for nonsmokers. These findings call into question the validity of treating responses from any one question about lung-cancer probability or frequency as reflections of stable, well-formed beliefs of the respondents. Even smokers are likely to hold only vague impressions of certainty about various smoking hazards, and their construction of a precise numeric estimate is highly malleable depending on the researcher's choice of question type and the context in which the question is posed.

GENERAL DISCUSSION AND CONCLUSION

I have argued that there are three substantial reasons to avoid comparing subjective and objective probability/frequency estimates for a single event in order to determine whether people have appropriate expectations about the possibility of the event. The first reason concerns problems with respondents misusing or misinterpreting the numeric probability scale. The second reason concerns the possibility that the numeric subjective probability measure would fail to assess the more intuitive and non-analytic components of certainty that mediated the decision or behavior that is of ultimate interest to the researcher. The third reason, supported by data from Experiments 1 and 2, concerns the fact that subjective probabilities often reflect ad hoc processes rather than stable beliefs and, therefore, can be highly susceptible to wording and context effects.

Although I am arguing against using comparisons between subjective and objective estimates to determine whether people are appropriately sensitive on the possibility of an event, this is importantly distinct from saying that researchers cannot draw conclusions about whether people overestimate or underestimate the numeric probability of a target event. The argument is also importantly distinct from saying that subjective probability estimates can never be used to assess people's expectations about the possibility of an event. There are in fact many conditions in which subjective probability estimates provide excellent information about people's expectations for a specified event. Below, I mention some key factors that can influence the appropriateness of drawing conclusions about internal expectations and representations of certainty from people's responses to subjective probability questions.

One factor that can improve one's ability to draw valid conclusions about internal expectations from subjective probability estimates for a target event is to examine the estimates for the event in the context of estimates for other events (for a related argument see Varey *et al.*, 1990). This is a built-in property of some studies (e.g. many studies focusing on the confidence-accuracy relation), but survey researchers and others can utilize this strategy as well when trying to interpret probability responses for a particular event. This strategy can help the researcher notice and possibly adjust for at least three concerns. First, if a sample of respondents has a general propensity to, for example, overestimate the likelihood of low-probability events, this fact will not be known if a researcher focuses only on one event and compares estimates for that event to an objective standard. If, however, the researcher solicits estimates for a focal event and several other events (regardless of whether the same subjects give estimates for all events or different subjects give estimates for different events), overall patterns of overestimation or underestimation can be used to help interpret or correct estimates for an individual event. Second, interpreting an estimate for one event in the context of other events can also help alleviate, although not eliminate, concerns over some contextual influences (such as those demonstrated in Experiments 1 and 2)—assuming those contextual influences have approximately equal impact on responses about all events. For example, if a particular question wording tends to artificially increase estimates for all events, an interpretation that focuses on the *relative* standing of estimates for the target event will be unaffected, but an interpretation of the estimate on an absolute scale (without respect to its relative standing) will be biased. Third, interpreting an estimate within the context of other estimates can partially alleviate problems that arise if some or all participants interpret the probability scale as reflecting concern. If the scale is treated as a scale of concern, then, as alluded to above, interpretation of a 'probability estimate' for any event in isolation would be just as tenuous as the interpretation of a Likert-type rating of concern. As with a Likert-type response, however, a researcher can better interpret the meaning of a probability/concern response by assessing how the response compared to responses for other events.

A second factor that can affect a researcher's ability to draw valid conclusions about internal expectations from a probability estimate is the degree to which the respondents have experience with the probability scales. The analogy mentioned earlier about basketball players estimating the distance between the basket and the three point line illustrated that inaccuracies in judgment can sometimes be accounted for by a lack of familiarity with a response unit rather than an inaccuracy in the internal representation that the researcher is investigating. The influence of bias in magnitude judgments decreases as familiarity of the response units increases (Poulton, 1989.) Relatedly, it seems appropriate to assume that as subjects' familiarity with the probability scale increases, their skill at mapping a set of internal beliefs about the likelihoods of events onto a numeric probability scale improves. Hence, if a group of weather forecasters (with extensive experience making probability judgments) and a random sample of Americans both provided a mean estimate of 45% for a smoker's chance of dying from lung cancer, it would seem more acceptable to claim that the weather forecasters have an inappropriately high expectation of the target event than it would to draw the same conclusion about the random sample of Americans.

A third factor that can affect a researcher's ability to draw valid conclusions about internal expectations from a probability estimate is related to the conceptual interests of the researcher. If a researcher is interested specifically in beliefs in objective probability (and has reason to ignore intuitive perceptions of certainty), and if he/she can safely assume that respondents have treated the probability scale as a true probability scale rather than an intuitive scale or a concern scale, then comparing subjective probabilities for an event to objective standards can be appropriate. For example, if a researcher is interested in participants' ability to determine the correct probability of an event through rule-based reasoning (e.g. What is the probability that all three coin flips will be heads?), then comparing a response to an absolute standard can be appropriate. As alluded to above, however, the researcher should be wary of assuming that a respondent's belief in objective probability is synonymous with his or her gut-level intuitions about the event, which are often best assessed with nonnumeric measures (Kirkpatrick and Epstein, 1992; Windschitl and Wells, 1996, 1998).

There are certainly other conditions under which comparisons between subjective and objective probability estimates for a given event are appropriate. The important point here is that whenever such comparisons are made, at least three questions, which relate to the three arguments described earlier, should be considered. First, did the respondents interpret the probability scale as such and not as a scale of concern or some related dimension, and were the respondents adequately familiar with how to use the scale? Second, did the numeric probability responses capture the internal representations of certainty that are of most interest to the researcher, or is it possible that the numeric responses reflect beliefs in objective probability, while the critical construct (i.e., the construct that actually mediates judgments and behaviors relevant to the target event) is a more intuitive representation of certainty? Third, given that probability responses reflect, in large part, ad hoc constructions of belief rather than beliefs retrieved from memory, did aspects of the survey format or the target question itself have a substantial impact on the numeric responses? Without adequately addressing these questions, subjective estimates for a single event, such as those discussed by Viscusi (1990, 1991, 1992), cannot lead to firm conclusions about the appropriateness of people's expectations about that event.

Finally, it is important to emphasize a point about the role of response type in light of the present arguments. Many studies have revealed important distinctions between frequency and probability estimates, and there is an ongoing discussion of whether certain types of bias may be less severe when frequentist rather than probabilistic representations are used (see e.g. Griffin and Buehler, 1999; Harries and Harvey, 2000; Howell and Burnett, 1978; Howell and Kerkar, 1982; Gigerenzer, 1994; Gigerenzer and Hoffrage, 1995, Slovic and Monahan, 1995; Slovic *et al.*, 2000). However, there is no reason to assume that frequency judgments are immune to the potential problems described throughout in this paper. In fact, in Experiments 1 and 2, frequency estimates showed the same type of sensitivity to the wording and context manipulations as did probability estimates. Hence, regardless of the response format used by a researcher, determining whether a respondent or group of respondents is adequately sensitive to the possibility of an event requires more than just a comparison between subjective numeric estimates and the relevant objective standard.

APPENDIX: THE EXPLICIT-REPRESENTATION VERSION OF THE PROBABILITY AND FREQUENCY QUESTIONS USED IN EXPERIMENT 1

What is the chance that you will die from each of the following? Place a chance estimate next to each cause of death listed below. Because the list includes all possible causes of death, and because a person's death is attributed to only one cause, *your responses should add up to 100%.*

______% household or workplace accident
______% car accident
______% natural disaster (e.g. flood, tornado, earthquake)
______% leukemia
______% lung cancer

______% liver or kidney disease
______% Alzheimer's disease or Lou Gehrig's disease
______% heart disease
______% stroke
______% all other possible causes
TOTAL = 100%

Among 100 people who smoke 2 packs per day, how many would die from each of the following? Because the list includes all possible causes of death, and because a person's death is attributed to only one cause, *your responses should add up to 100*.

______ household or workplace accident
______ car accident
______ natural disaster (e.g. flood, tornado, earthquake)
______ leukemia
______ lung cancer
______ liver or kidney disease
______ Alzheimer's disease or Lou Gehrig's disease
______ heart disease
______ stroke
______ all other possible causes
TOTAL = 100

REFERENCES

Black, W.C., Nease, R.F. Jr, & Tosteson, A.N.A. (1995). Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. *Journal of the National Cancer Institute*, **87**, 720–731.

Borland, R. (1997). What do people's estimates of smoking related risk mean? *Psychology and Health*, **12**, 513–521.

Bruine de Bruin, W., Fischhoff, B., Millstein, S.G., & Halpern-Felsher, B.L. (in press). Verbal and numerical expressions of probability: 'It's a fifty-fifty chance.' *Organizational Behavior and Human Decision Processes*.

Budescu, D.V., & Wallsten, T.S. (1995). Processing linguistic probabilities: general principles and empirical evidence. *The Psychology of Learning and Motivation*, **32**, 275–317.

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: when people behave against their better judgment. *Journal of Personality and Social Psychology*, **66**, 819–829.

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

Eiser, J.R. (1994). Risk judgements reflect belief strength, not bias. *Psychology and Health*, **9**, 197–199.

Eiser, J.R., & Hoepfner, F. (1991). Accidents, disease, and the greenhouse effect: effects of response categories on estimates of risk. *Basic and Applied Social Psychology*, **12**, 195–210.

Epstein, S. (1990). Cognitive-Experiential Self-Theory. In L. Pervin (Ed.), *Handbook of personality: Theory and Research* (pp. 165–192). New York: Guilford.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist*, **49**, 709–724.

Fischhoff, B., & Bruine de Bruin, W. (1999). Fifty-fifty = 50%? *Journal of Behavioral Decision Making*, **12**, 149–163.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault Trees: sensitivity of estimated failure probability to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, **4**, 330-344.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In G. Wright & P. Ayton (Eds.), *Subjective Probability* (pp. 130–161). Chichester: John Wiley.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review*, **102**, 684–704.

Gonzalez, M., & Frenck-Mestre, C. (1993). Determinants of numerical versus verbal probabilities. *Acta Psychologica*, **83**, 33–51.

Griffin, D., & Buehler, R. (1999). Frequency, probability, and prediction: Easy solutions to cognitive illusions? *Cognitive Psychology*, **38**, 48–78.

Hanson, J.D., & Kysar, D.A. (1999). Taking behavioralism seriously: some evidence of market manipulation. *Harvard Law Review*, **112**, 1422–1572.

Harries, C., & Harvey, N. (2000). Are absolute frequencies, relative frequencies, or both effective in reducing cognitive biases? *Journal of Behavioral Decision Making*, **13**, 431–444.

Howell, W.C., & Burnett, S.A. (1978). Uncertainty measurement: a cognitive taxonomy. *Organizational Behavior and Human Performance*, **22**, 45–68.

Howell, W.C., & Kerkar, S.P. (1982). A test of task influences in uncertainty measurement. *Organizational Behavior and Human Performance*, **30**, 365–390.

Kirkpatrick, L.A., & Epstein, S. (1992). Cognitive-experiential self-theory and subjective probability: further evidence for two conceptual systems. *Journal of Personality and Social Psychology*, **63**, 534–544.

Klein, W.M. (1997). Objective standards are not enough: Affective, self-evaluative, and behavioral responses to social comparison information. *Journal of Personality and Social Psychology*, **72**, 763–774.

Ofir, C., & Reddy, S.K. (1996). Measurement errors in probability judgements. *Management Science*, **42**, 1308–1325.

Pacini, R., & Epstein, S. (1999). The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, **76**, 972–987.

Poulton, E.C. (1989). *Bias in quantifying judgments*. Hove: Lawrence Erlbaum Associates.

Robinson, L.B., & Hastie, R. (1985). Revision of beliefs when a hypothesis is eliminated from consideration. *Journal of Experimental Psychology: Human Perception and Performance*, **11**, 443–456.

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: advances in support theory. *Psychological Review*, **104**, 406–415.

Schwartz, N., Hippler, H.J., Deutsch, B., & Strack, F. (1985). Response scales: effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, **49**, 388–395.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, **119**, 3–22.

Slovic, P. (2000a). What does it mean to know a cumulative risk? Adolescents' perceptions of short-term and long-term consequences of smoking. *Journal of Behavioral Decision Making*, **13**, 259–266.

Slovic, P. (2000b). Rejoinder: the perils of Viscusi's analyses of smoking risk perceptions. *Journal of Behavioral Decision Making*, **13**, 273–276.

Slovic, P. (in press). Rational actors and rational fools: The influence of affect on judgment and decision making. *Roger Williams University Law Review.*

Slovic, P., & Monahan, J. (1995). Probability, danger and coercion: a study of risk perception and decision making in mental health law. *Law and Human Behavior*, **19**, 49–65.

Slovic, P., Monahan, J., MacGregor, D.G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, **24**, 271–295.

Smith, E.R., & DeCoster, J. (1999). Associative and rule-based processing: A connectionist interpretation of dual-process models. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology.* (pp. 323–336). New York: The Guilford Press.

Smith, E.R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, **4**, 108–131.

Teigen, K.H. (1974). Subjective sampling distributions and the additivity of estimates. *Scandinavian Journal of Psychology*, **15**, 50–55.

Teigen, K.H. (1983). Studies in subjective probability III: the unimportance of alternatives. *Scandinavian Journal of Psychology*, **24**, 97–105.

Teigen, K.H. (in press). When equal chances = good chances: verbal probabilities and the equiprobability effect. *Organizational Behavior and Human Decision Processes*.

Teigen, K.H., & Brun, W. (2000). Ambiguous probabilities: when does $p = .3$ reflect a possibility, and when does it express a doubt? *Journal of Behavioral Decision Making*, **13**, 345–362.

Tversky, A. & Koehler, D.J. (1994). Support theory: a nonextensional representation of subjective probability. *Psychological Review*, **101**, 547–567.

Varey, C.A., Mellers, B.A., & Birnbaum, M.H. (1990). Judgments of proportions. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 613–625.

Viscusi, W.K. (1990). Do smokers underestimate risks? *Journal of Political Economy*, **98**, 1253–1269.

Viscusi, W.K. (1991). Age variations in risk perceptions and smoking decisions. *The Review of Economics and Statistics*, **4**, 577–588.

Viscusi, W.K. (1992). *Smoking: Making the risky decision*. New York: Oxford University Press.
Viscusi, W.K. (2000). Comment: the perils of qualitative smoking risk measures. *Journal of Behavioral Decision Making*, **13**, 267–271.
Wallsten, T.S., Budescu, D.V., & Zwick, R. (1993). Comparing the calibration and coherence of numerical and verbal probability judgments. *Management Science*, **39**, 176–190.
Weinstein, N.D. (1998). Accuracy of smokers' risk perceptions. *Annals of Behavioral Medicine*, **20**, 135–140.
Windschitl, P.D. (2000). The binary additivity of subjective probability does not indicate the binary complementarity of perceived certainty. *Organizational Behavior and Human Decision Processes*, **81**, 195–225.
Windschitl, P.D., Martin, R., & Flugstad, A.R. (in press). Context and the interpretation of likelihood information: The role of intergroup comparisons on perceived vulnerability. *Journal of Personality and Social Psychology.*
Windschitl, P.D., & Weber, E.U. (1999). The interpretation of 'likely' depends on the context, but '70%' is 70%—right? The influence of associative processes on perceived certainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 1514–1533.
Windschitl, P.D., & Wells, G.L. (1996). Measuring psychological uncertainty: verbal versus numeric methods. *Journal of Experimental Psychology: Applied*, **2**, 343–364.
Windschitl, P.D., & Wells, G.L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology*, **75**, 1411–1423.
Windschitl, P.D., & Young, M.E. (in press). The influence of alternative outcomes on gut-level perceptions of certainty. *Organizational Behavior and Human Decision Processes*.
Wright, G., & Walley, P. (1983). The supra-additivity of subjective probability. In B.P. Stigum & F. Wenstøp (Eds.), *Foundations of utility and risk theory with applications* (pp. 233–244). Dordrecht: Reidel.
Yamagishi, K. (1997). When a 12.86% mortality is more dangerous than 24.14%: implications for risk communication. *Applied Cognitive Psychology*, **11**, 495–506.
Zimmer, A.C. (1983). Verbal vs. numeric processing of subjective probabilities. In R. Scholz (Ed.), *Decision making under uncertainty* (pp. 159–182). Amsterdam: North-Holland.

Author's biography:
Paul Windschitl is an Assistant Professor of Psychology at the University of Iowa. His research interests include perceptions of certainty, subjective probability judgments, and causal attributions.

Author's address:
Paul D. Windschitl, Department of Psychology, University of Iowa, Iowa City, Iowa, 52242, USA.

