

The Interpretation of “Likely” Depends on the Context, but “70%” Is 70%—Right? The Influence of Associative Processes on Perceived Certainty

Paul D. Windschitl
University of Iowa

Elke U. Weber
Columbia University

Past research has demonstrated that interpretations of vague verbal forecasts (e.g., “likely”) differ as a function of the context to which they refer. Experiments 1 and 2 demonstrate that precise numeric forecasts (e.g., “70%”) are also susceptible to such context effects. Participants read descriptions of target events and experts’ numeric forecasts. Perceptions of certainty, expressed on nonnumeric scales, differed as a function of context manipulations. The results of Experiments 3a, 3b, and 4 indicate that these effects can be mediated by perceptions of an event’s representativeness independently of subjective base rates. The results are also consistent with the idea that two types of semi-independent processing—associative and rule based—can have important influences on perceptions of certainty. Implications of this distinction for research on judgments and decisions under uncertainty are discussed.

Imagine that you are coordinating the installation of new communication systems at European sites and that some of the installation work requires perfectly rainless weather. You meet with a reputable meteorological expert hired by your company to discuss the day’s forecasts for various work sites. The meteorological expert tells you there is “a slight chance of rain” for the London site. Worried about the possibility of rain, you decide that it is probably not a good idea to schedule critical work at the London site. Later in your meeting, the expert tells you there is “a slight chance of rain” for the Madrid site. With little concern about rain, you decide that it is okay to schedule critical work at that site. The expert forecasted “a slight chance of rain” for both sites, yet differing perceptions about the chances of rain led you to different decisions about the two sites.

Research about the effects of context on the interpretations of uncertainty phrases suggests that this is a plausible scenario (see Brun & Teigen, 1988; Wallsten, Fillenbaum, & Cox, 1986; Weber & Hilton, 1990). This research has demonstrated that the numeric interpretations associated with certainty phrases such as “quite likely” or “good chance” vary as a function of the events that these phrases qualify. Thus, people provide different numeric interpretations of the phrase “slight chance” when it refers to the probability that an ankle is sprained than when it refers to the probability of suffering life-threatening side effects from a flu shot (see Wallsten et al., 1986; Weber & Hilton, 1990). Similarly, the phrase “a slight chance of rain” will have one

interpretation in reference to London and another in reference to Madrid.

Now imagine that instead of having provided certainty phrases as forecasts, the meteorological expert in the above scenario provided precise numeric probabilities as forecasts. If the expert stated that there was a 5% chance of rain in both London and Madrid, is it nevertheless possible that your uncertainty regarding the likelihood of rain might differ for the two work sites?

On the basis of previous accounts of how context information influences the interpretations of verbal certainty phrases, one would not expect that context should also influence the interpretations of numeric probability forecasts. The predominant account, which we’ll call the *base-rate account*, assumes that verbal certainty phrases have a range of plausible interpretations and that contextual factors influence the selection of an appropriate numeric translation (see Wallsten et al., 1986; Weber, 1994; Weber & Hilton, 1990). More specifically, the account proposes that the interpretation of a verbal certainty phrase is a weighted average of two vague probabilities—one reflecting the perceived meaning of the phrase in isolation and one reflecting the interpreter’s subjective probability for the event within the given context (i.e., without any forecast information). This account has received support in research in which participants interpreted the meaning of verbal forecasts when used to describe events with high base rates (e.g., “likely” to snow in North Carolina mountains in December) and low base rates (e.g., “likely” to snow in North Carolina mountains in October). As predicted, the same verbal forecasts were interpreted as indicating higher probabilities when they referred to a high base-rate event than when they referred to a low base-rate event (Wallsten et al., 1986). Another account for context effects suggests that, in addition to base rates, the severity of an outcome might play a role. Weber and Hilton (1990) found that a certainty

Paul D. Windschitl, Department of Psychology, University of Iowa; Elke U. Weber, Department of Psychology, Columbia University.

Correspondence concerning this article should be addressed to Paul D. Windschitl, Department of Psychology, University of Iowa, Iowa City, Iowa 52242. Electronic mail may be sent to paul-windschitl@uiowa.edu.

phrase was given a higher numeric interpretation when it referred to a severe outcome than when it referred to a less severe outcome (see also Weber, 1994). Although both of these accounts may be valid descriptions of many types of context effects, both appear to be limited to situations in which the forecast information is vague rather than precise.

In this article, we propose an account of how context might influence perceptions of certainty even when precise numeric forecast information is available. Our *representativeness account* assumes that the strengths of mental associations between a specified context and event can influence perceived certainty without mediation through subjective probability estimates. Rather, the strength of mental associations can affect the perceived representativeness of the event for a given context, and representativeness can have a direct effect on perceptions of whether the event will occur. According to this account, even when a precise numeric forecast is believed to be the best estimate of an event's probability, this belief does not preclude an effect of representativeness. For example, a person can feel more optimism about a rainless day in Madrid versus London even though the person believes there is a 5% chance of rain in both places.

The hypothesis that associative processes can influence perceptions of certainty without affecting beliefs in objective probability shares important similarities to a broader distinction between two types of information processing. Numerous cognitive and social theorists have suggested a variety of dichotomies relevant to information processing (e.g., Bruner, 1986; Epstein, Lipson, Holstein, & Huh, 1992; Langer, 1989; Schneider & Shiffrin, 1977; see also Chaiken & Trope, 1999). Sloman's (1996) distinction between rule-based and associative processing is most relevant for the present work. According to Sloman (1996), rule-based processing is a relatively controlled form of processing that operates according to formal rules of logic and evidence, is governed by hard constraints, and is mediated by conscious appraisals of information and events. Associative processing is a more spontaneous form of processing that operates by principles of similarity and temporal contiguity, is governed by soft constraints, and is not mediated by conscious appraisals.

Among these differences between associative and rule-based processing, the distinction that we consider most relevant for the present work concerns the issue of rule execution. A response driven by rule-based processing follows from the execution of one or more rules that are assumed to be relevant to the task (e.g., mathematical rules, *modus ponens*, the conjunction rule). Executing those rules requires that a respondent first represent the information in a form that is compatible with the rules and then manipulate the information according to his or her understanding of the rules. In strictly associative processing, responses are not mediated by the execution of rules. Rather, concept activation influences responses directly, just as associatively based priming influences the recognition of a target word. Pathways and patterns of activation follow principles of similarity and temporal contiguity; the stronger the association between two concepts (which depends on similarity, re-

peated exposure, etc.), the more activation will pass from one to another.

According to Sloman (1996), associative and rule-based processing typically work in concert to guide reasoning and decision making. However, a key point for the present work is that these two forms of processing are semi-independent. In support of this proposal, Sloman provided examples from reasoning, categorization, and judgment research in which people find two simultaneously contradictory responses—one presumably mediated by associative processing and the other by rule-based processing—to be compelling for a given problem. For example, although people know that a whale does not fit the classification of "fish," common phrases like "technically a whale is a mammal" suggest that people are influenced by the similarity between whales and fish (or dissimilarity between whales and other mammals). Also, even after people fully understand how to apply the conjunction rule to the Linda problem, they nevertheless maintain a nagging feeling that Linda is more likely to be a bank teller active in the feminist movement than a bank teller (Sloman, 1996).

Along similar lines, the representativeness account that we are proposing for context effects assumes that a person can hold a belief about the objective probability of an event yet can also be influenced by event-context associations. Specifically, strong event-context associations can underlie perceptions of representativeness, which influence certainty independently of a forecaster's estimate. Hence, the representativeness account predicts context effects even when an expert's forecast is a precise numeric probability (tested in Experiments 1 and 2). We contrast this account with the base-rate account (see Experiments 3a, 3b, and 4). Previous descriptions of the base-rate account have not specified whether the weighted averaging of the relevant subjective probabilities is a part of a deliberate and conscious strategy or a more automatic and preconscious process. It is also possible to interpret the base-rate account as a functional model of context effects without making any claims about the processes mediating the effect. We start by interpreting the account as a rule-based description of context effects (where the critical mediator of a likelihood judgment is a weighted averaging process), but our experiments also address the question of whether a functional base-rate account can explain the effects observed here.¹

Measuring Associative-Based Influences on Certainty

One difficulty in testing the prediction of the representativeness account (i.e., that context will affect certainty even when an expert's forecast is a precise probability)

¹ The base-rate account does not preclude the involvement of associative processes. For example, one could assume that basic associative memory processes are used to generate a subjective base-rate estimate. Nevertheless, the base-rate account could be considered rule based given the assumption that some execution of a weighted averaging process (which serves to incorporate the subjective base-rate estimate) is the critical mediator of people's probability responses.

concerns the issue of how associative-based influences on certainty should be measured. Given that we are interested in how context influences certainty when a forecast is numeric, using numeric probability scales to measure participants' certainty would be problematic. Responses on such a measure can map directly onto the forecast information; if we gave participants context information and a numeric forecast about an event and then asked "How likely is Event X?" a respondent could simply restate the forecast that was given. Recent research investigating novel influences on perceived certainty has successfully used alternative measures of subjective certainty (Windschitl & Wells, 1996, 1998). These measures ask respondents to indicate their certainty by selecting a response from an ordered list of verbal phrases such as "quite likely" and "almost impossible." Windschitl and Wells (1996) compared the effectiveness of verbal versus numeric measures in detecting subtle variations in perceived certainty—variations that were not likely to be functions of rule-based considerations on the part of respondents. Verbal measures were more sensitive than numeric measures to these variations. For example, some participants read a scenario in which a woman was drawing from a box containing 1 winning ticket out of 10 total; other participants read that the box contained 100 winning tickets out of 1,000 total. When asked how likely it was that the woman would draw a winning ticket, participants' responses on a numeric probability scale were not sensitive to this manipulation, which in other research has been shown to affect several judgments and behaviors mediated by perceived certainty (Denes-Raj, Epstein, & Cole, 1995; Kirkpatrick & Epstein, 1992; Miller, Turnbull, & McFarland, 1989). However, responses on a verbal certainty measure did show that participants were more optimistic in the 100-in-1,000 case than the 1-in-10 case. In another experiment, responses on verbal rather than numeric measures of certainty were found to be better predictors of how participants reported they would behave in a variety of scenarios involving uncertain events (Windschitl & Wells, 1996, Experiment 3).

Windschitl and Wells (1998) also used verbal measures to detect and investigate the alternative-outcomes effect. For example, participants read a scenario in which they held 21 raffle tickets and the 5 remaining players held 14, 13, 15, 12, and 13 tickets or 52, 6, 2, 2, and 5 tickets. This manipulation, which does not vary the objective likelihood of the participant winning, did not have a significant effect on participants' numeric estimates of their winning. However, a verbal measure of perceived certainty detected a robust alternative-outcomes effect showing that people feel less certain about winning when there is an alternative outcome that is more likely than the target outcome (i.e., their winning). Additional experiments confirmed that manipulations of this type influence not only verbal estimates of certainty but also other relevant judgments and choice behaviors (Windschitl & Wells, 1998).

Experiment 1

Given the success with which verbal measures have been used in recent research to investigate novel influences on

perceived certainty, we used a verbal measure to test whether context might influence perceptions of certainty even when a precise and credible numeric forecast is known. Participants in Experiment 1 read scenarios describing events with unknown outcomes. Each scenario included a precise probability forecast from a knowledgeable expert (e.g., doctor, executive planning director) as well as information about the context. We manipulated information in a between-subjects fashion; the London–Madrid manipulation that we discussed above is typical of these manipulations. After reading the scenario, participants indicated their perceived certainty in the target event on a verbal measure. Because we propose that context effects can be caused by differences in representativeness of event–context pairs, and that the processes mediating such effects are semi-independent from beliefs in objective probability, we expected the context manipulations to produce differences in verbal certainty responses even though the numeric forecast information was complete and precise. Finally, we asked participants to indicate, from memory, the numeric estimates of likelihood given by the experts.

Method

Participants. The participants were 92 undergraduate students at The Ohio State University who received credit in an introductory psychology course.

Materials. We constructed two versions (A and B) of seven scenarios for the experiment. Each scenario described a focal event for which an expert had provided a precise probabilistic estimate. Context was manipulated between the two versions of each scenario in a between-subjects fashion. We designed the versions such that the focal event would appear more representative of the context described in one version than in the other version.² Appendix A contains summary information regarding the versions of the seven scenarios used in the experiment. Versions A of the scenarios were grouped into one scenario packet, and Versions B were grouped into another. Scenarios 3 and 7 read as follows, with the parenthetical information varying between the two versions (A/B):

Carol is a premed student who is very excited about being a physician. She has finished Calculus 1 and 2 and received [Cs/Bs] in the two classes. One of the requirements for premed students is to complete either the sequence Calculus 1, Calculus 2, Calculus 3, or the sequence Calculus 1, Calculus 2, Research Methods. Her academic counselor tells her that school records indicate that students who received [Cs/Bs] in the two previous classes have a seventy percent chance of passing Calculus 3. Carol has to decide whether to take Calculus 3, given that she got [Cs/Bs] in the two previous courses, or Research Methods. You have been taking classes

² The context manipulations are discussed in terms of representativeness, yet one might note that some manipulations appear to vary the availability (or accessibility) of the events. We think that the nature of the experimental materials blurs the distinction between these two concepts. Events that seem very representative of a context are also likely to be mentally available given the contextual cue. For example, rain seems very representative of London, and rainy images are easily accessible when cued by the London context. We do not attempt to disentangle these concepts; we assume that both are mediated by associative processes.

with her for the past 2 years and know that she gets [Cs/Bs], yet takes her medical and academic career very seriously.

Janet is planning to go on a year-long trip to [Hawaii/India], where she will take a teaching position [close to Waikiki beach/in Calcutta]. She is looking forward to combining work with play. Before leaving, her doctor gives her an extensive physical. During the physical, Janet finds out that she has a common blood condition that makes her more susceptible to certain ailments. Her doctor tells her that while in [Hawaii/India], there is a thirty-percent chance that she will contract a mild form of malaria. She is leaving for [Honolulu/Calcutta via Bombay] in a couple of weeks, and since she is your best friend, you will be sad not to have her around for the year.

After reading each scenario, respondents were asked to indicate how likely they perceived the focal event to be. For example, the uncertainty question for Version A of Scenario 7 was "Please mark on the rating scale below how likely you think it is that Janet will contract malaria while in Hawaii." The rating scale that we used, shown in Figure 1, is an adaptation of the verbal scales introduced by Windschitl and Wells (1996). The verbal expressions that accompanied this scale were the same as those used on the 11-point, discrete, response scale that appeared in their second experiment.

We also constructed recollection questionnaires for the experiment. The questionnaires contained seven questions, one for each scenario, designed to test participants' memories of the experts' numeric predictions. Each question asked respondents to indicate the percentage estimate for the focal event given by the expert. About half of the participants were asked to provide their responses by circling 1 of 21 percentage estimates (0%, 5%, 10%, etc.), and half were asked to generate a percentage response. This response-format variation had no effect on participants' responses and is not discussed further.

Procedure. We tested participants in groups ranging in size from 4 to 10. They read and provided uncertainty responses for each scenario in their scenario packet. We collected the packet from them on completion. After a short delay period (either 1, 25, or 45 min), participants were given the recollection questionnaire, and they completed it at their own pace. Length of the delay did not affect participants' responses and is not discussed further.

Results and Discussion

We scored participants' uncertainty responses (made on the scale shown in Figure 1) in millimeters, and scores could vary from 0 to 150. Participants' perceptions of uncertainty clearly were sensitive to the numeric predictions provided by experts in the scenarios. The rank ordering of the scenarios based on the experts' numeric forecasts (Scenario 5 with its 20% forecast ranked lowest and Scenario 3 with its 70% forecast ranked highest) was nearly identical to the rank ordering based on participants' mean responses; the only disagreement in the two orderings involved the ranks for Scenarios 2 and 4.³

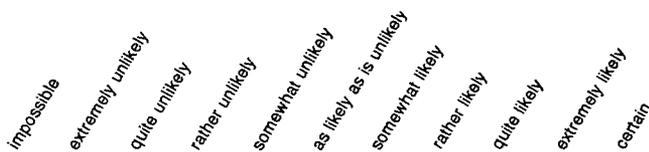


Figure 1. The verbal response scale that accompanied the uncertainty questions in Experiment 1.

Table 1
Mean Uncertainty Responses and Effect-Size Estimates
for Scenarios of Experiment 1

Scenario	Version				<i>d</i>
	A		B		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1	77.9	32.1	85.4	26.3	0.25
2	74.4	26.9	89.7	27.8	0.56
3	93.1	25.5	118.3	20.5	1.06
4	69.5	24.1	91.7	18.7	1.00
5	66.2	30.6	60.5	35.6	0.18
6	81.3	22.6	110.2	19.1	1.36
7	60.9	24.2	71.7	25.8	0.44

Note. The uncertainty responses were scored in millimeters; the maximum score was 150. The *d* column displays the standardized mean difference in responses between the A and B versions. For all scenarios except No. 5, Version B was the high-representativeness version. In Scenario 5, Version A was the high-representativeness version.

To determine whether the contextual factors that were manipulated between versions of the scenario had a significant effect on perceived certainty, we submitted participants' responses to a multivariate analysis of variance (MANOVA). As predicted, the analysis showed that participants expressed significantly more certainty for scenario versions that contained contexts that were designed as high in representativeness versus those designed as low in representativeness, $F(7, 84) = 9.82, p < .001$.⁴ Table 1 displays means and standard deviations for the two versions of each scenario, as well as effect-size estimates for the differences between means. The average effect size across the seven scenarios was 0.69 standard deviations, a medium-to-large effect.⁵ The average difference in responses between the two versions of the scenarios was about 15 mm, which is exactly the distance separating two adjacent verbal labels as they appeared on the uncertainty scale. These results clearly indicate that the context manipulations had substantial effects on participants' perceptions of uncertainty.

To better understand the nature of the context effects, consider Scenarios 3 and 7 that were shown earlier. For

³ Although some readers might wish to judge the agreement between the experts' forecasts and participants' responses by scoring responses as percentages (with, for example, a 75-mm response scored as 50%), we caution that there is no basis nor necessity to assume that responses on this verbal measurement device should be mapped directly into numeric probabilities (for further discussion of this issue, see Windschitl & Wells, 1996).

⁴ Conducting this analysis required that we change the signs (i.e., + and -) of the responses for Scenario 5, because a given participant was not presented with exclusively high-representativeness or low-representativeness contexts. This transformation does not affect the interpretation of the overall analysis, but it does render the transformed means that are collapsed across scenarios largely uninterpretable. Therefore, we present the untransformed means separately for each scenario in Table 1.

⁵ Cohen (1988) suggested that effects with magnitudes of .20, .50, and .80 should be considered "small," "medium," and "large" effects, respectively.

Scenario 3, some participants were told that Carol had received Cs in Calculus 1 and 2 and that, based on school records, students with such a record have a 70% chance of passing Calculus 3. The mean uncertainty estimate provided by these participants was located near the "somewhat likely" label (93 mm), whereas the mean uncertainty estimate for participants who received the alternative version of the scenario, in which she received Bs, was located near the "quite likely" label (118 mm). For Scenario 7, some participants were told that Janet was taking a trip to Hawaii and that, because of a blood condition, she had a 30% chance of contracting a mild form of malaria on her trip to Hawaii. The mean uncertainty estimate of these participants was located at the "somewhat unlikely" label (61 mm), whereas the mean uncertainty estimate of the participants who received the alternative version, in which Janet is going to India, was located near the "as likely as unlikely" label (72 mm). As stated earlier, we designed the two versions of each scenario such that the focal event would appear more representative of the context in one version than in the other version. For example, we assumed that in the minds of our participants, contracting malaria is more representative of a visit to India than to Hawaii, and passing Calculus 3 is more representative of "B" students than of "C" students.

It is important to note that the numeric forecast provided in the scenario takes into consideration the context. For example, the doctor's prediction regarding Janet's chances of contracting a form of malaria are specific to the location she will be visiting. Therefore, there is no normative reason for participants to adjust their uncertainty based on her trip destination. Our data indicate, however, that people do allow context to influence their uncertainty even in the presence of a fully relevant, credible, and precise numeric forecast.

One alternative explanation for these data would be that a significant proportion of people did not fully encode the numeric likelihood information about the focal event of a given scenario; their responses, consequently, would be based solely on their own perceptions of how likely the event would be in the described context. Responses on the recollection questionnaire ruled out this explanation. If the explanation were correct, one would expect that those participants who failed to encode the numeric information and based their uncertainty responses solely on contextual information would also need to base their memory reports of the experts' likelihood estimates on the contextual information. In other words, if a significant portion of participants failed to encode the experts' numeric forecasts, there should be evidence that the context manipulations influenced responses on the memory questionnaires. Yet, no such evidence was found. Context did not have a significant effect on the recall responses, $F(7, 79) < 1$. Table 2 displays the mean responses on the recall questions for each scenario version. In additional analyses on the verbal uncertainty responses, we excluded data points for which corresponding recall responses were incorrect (i.e., whenever a participant incorrectly recalled the forecasted probability, we excluded his or her uncertainty estimate from the analysis). The strengths of the context effects observed in these analyses were equivalent to those observed in the overall analyses. The mean of

Table 2
Mean Memory Responses for Scenarios of Experiment 1

Scenario	Version			
	A		B	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
1	55.8	16.4	52.8	13.5
2	40.6	14.1	38.7	13.8
3	67.1	10.4	65.5	11.4
4	48.6	15.6	49.6	15.3
5	30.0	16.3	31.9	22.6
6	55.9	13.5	59.4	11.3
7	33.3	07.3	32.6	10.3

the effect sizes across the scenarios was 0.71 (compared with 0.69 in the overall analysis). These findings preclude the possibility that the context effects of Experiment 1 are attributable to participants who did not adequately process the relevant numeric forecasts of the experts.⁶ We argue that a better explanation of the context effect assumes that participants were fully aware of the relevance of an expert's forecast and that they accepted the forecast as an appropriate numeric probability estimate for the target event. The perceived representativeness of the events in the manipulated contexts influenced perceptions of certainty independently of beliefs in objective probability.

Experiment 2

To what extent were the context effects demonstrated in Experiment 1 particular to the type of response scale that was used or to the scenarios that were constructed? We assume that it was not the verbal qualities of the response scale used in Experiment 1 that allowed the detection of context effects but rather the fact that the nonnumeric scale did not allow people to simply restate the experts' estimates. Hence, other types of scales might be sensitive to context effects like those demonstrated in Experiment 1. Experiment 2 tested for context effects with a graphic line scale that had only two anchors at the endpoints. This type of scale seems to be growing in use, and Russell and Bobko (1992) provided convincing evidence that such scales can hold advantages over the coarser Likert-type scales. Experiment 2 also included a new set of three scenarios in which the target events were described as having relevance to the reader rather than to another individual. We borrowed two of

⁶ Some readers might be interested in participants' overall accuracy in recalling the numeric estimates provided by the experts. Across all scenarios and participants, 67.8% of participants' recollections were within 5 percentage points of the experts' estimates. For Scenario 7, for example, 82.6% of participants' recollections were between 25 and 35%. This recall performance was not moderated by the length of the delay period between participants' exposure to the scenarios and their introduction to the recollection questionnaire. The mean number of correct responses per participant (with only an exact recollection considered correct) out of a total of 7 was 3.9 for the immediate group, 4.0 for the 25-min group, and 3.6 for the 45-min group, $F < 1$.

these scenarios from previous research that showed that interpretations of verbal forecasts differ as a function of context (Wallsten et al., 1986). We expected these same scenarios to produce significant context effects even though the vague verbal forecasts were replaced with precise numeric forecasts.

Method

Participants. The participants were 86 undergraduate students at Iowa State University who received extra credit points in a psychology course.

Materials and procedure. We used two versions of three scenarios in the experiment. As in Experiment 1, an expert's numeric forecast was presented in each scenario. This numeric forecast was identical across the two versions of the scenarios, but the context was manipulated between versions.

All of the scenarios asked the reader to imagine receiving expert advice from a medical doctor. Scenario 1 was similar to the Hawaii-India scenario (No. 7) of Experiment 1. Readers were asked to imagine that they were going on a trip to Hawaii (or India) and that their doctor told them that, given their blood condition, they had a 20% chance of contracting a disease related to Malaria on the trip. Scenarios 2 and 3 were adaptations of those used by Wallsten et al. (1986, see their Table 1) in their demonstration that the interpretations of verbal uncertainty phrases can differ as a function of context (see also Weber & Hilton, 1990). For these scenarios, we used nearly the same text, but we replaced the verbal forecasts with precise probability estimates. In Version A of Scenario 2, a doctor stated that "if you eliminate caffeine there is a 40% chance that your gastric disturbances will stop." In Version B, the doctor stated that "there is a 40% chance it [a wart] will grow back again within 3 months." In Version A of Scenario 3, a doctor stated that "there is a 3% chance of severe, life-threatening side effects [for a flu shot]." In Version B, the doctor stated that "there is a 3% chance that it [your ankle] is sprained rather than broken."⁷

Participants read one version of each scenario. After each scenario they were asked about their uncertainty, and they provided a response by marking a line that was anchored at the endpoints. For Scenario 1, participants were asked "If you did go to Hawaii [or India], would you get the disease?" The response line was anchored on the left by *definitely wouldn't get the disease* and on the right by *definitely would get the disease*. For Version A and B of Scenario 2, the respective questions were "Will the gastric disturbances stop if you eliminate caffeine intake?" and "Will it [the wart] grow back?" For Versions A and B of Scenario 3, the questions were "Will you have severe side effects?" and "Is your ankle sprained?" Appropriate anchors (e.g., *definitely won't [will] have side effects*) appeared on the left and right ends of the response lines.

Results and Discussion

As expected, an overall MANOVA revealed a robust context effect across the scenarios, $F(3, 82) = 14.43, p < .001$. Participants' responses were scored from 0 to 150 mm. For Scenario 1, participants who read that they were going to Hawaii felt it was less likely that they would get a disease related to malaria ($M = 47.6, SD = 24.0$) than did participants who read that they were going to India ($M = 64.2, SD = 29.9$), even though participants in both groups were informed that their blood condition gave them a 20% of contracting such a disease in the location to which they were

traveling, $t(84) = 2.84, p < .01, d = 0.61$. For Scenario 2, the mean certainty response of participants reading about a 40% chance that a wart would grow back was 72.8 ($SD = 22.3$), whereas the mean response of participants reading about a 40% chance their gastric disturbances would stop was 88.8 ($SD = 27.9$), $t(84) = 2.94, p < .01, d = 0.63$. For Scenario 3, the mean certainty response of participants told there was 3% chance of developing a severe side effect was 25.9 ($SD = 25.7$), whereas the mean response of participants told there was a 3% chance of having a sprained ankle was 68.5 ($SD = 47.6$), $t(84) = 5.16, p < .001, d = 1.11$.

Although we did not directly assess the perceived representativeness between the events and contexts used in the three scenarios, we believe that differences in representativeness played a key role in producing the observed context effects. For example, we assume that, for our participants, malaria is more representative of diseases in India than of diseases in Hawaii (Scenario 1) and that an ankle sprain is more representative of a soccer injury than a life-threatening side effect is representative of a flu-shot reaction (Scenario 2). This explanation of the observed context effects is distinct from the base-rate account offered for previous research in which Scenarios 2 and 3 were used to test whether context influences people's numeric interpretations of verbal forecasts (Wallsten et al., 1986). The base-rate account assumes that context manipulations influence interpretations of verbal forecasts because the subjective base rates for the events differ across the manipulated contexts; the vagueness inherent in the verbal forecasts is resolved through a weighted averaging of the subjective base rates and the meaning of the verbal forecast itself. This explanation is a plausible account of how people might use subjective base rates to help interpret an expert's vague forecast.

In the present experiment, however, the forecasts were precise rather than vague. Participants had no basis for using their own perceptions of base rates to interpret the forecast or to make adjustments to the forecast. Given the wording of Scenario 3, for example, why would a participant decide that the doctor's probabilistic judgment needs adjustment? The judgment was precise and specific to the context, and the participant had no information about the symptoms on which the judgment was based. It does not appear that the base-rate account used to explain previous context effects can readily extend to account for the context effects observed here. Our representativeness account does not assume that participants adjusted their interpretation of the forecaster's estimate toward their perception of the base rate.

⁷ For nearly all of the scenarios used in the experiments reported here, *context manipulations* could be defined as manipulations that vary information that is related to the event but hold the general target event (e.g., rain, contracting malaria) constant. Two clear exceptions to this definition are the context manipulations in Scenarios 2 and 3 of Experiment 2. These manipulations, which we borrowed from Wallsten et al. (1986), require a broader definition for *context manipulation*. The target event itself was manipulated in these scenarios.

Rather, the perceived representativeness of an event for a given context influenced participants' perceptions of certainty independently of their beliefs in the forecaster's numeric estimate.

Although we argue that representativeness played a key role in the observed effects, factors other than representativeness may have also influenced the magnitude of the effects. Differences in the severity of the medical conditions and differences in asymmetric loss functions associated with the diagnostic errors may have influenced the context effects in Scenarios 2 and 3, although the direction of these influences was not necessarily the same as that of the representativeness manipulation (see Weber, 1994; Weber & Hilton, 1990). Another influence possibly underlying the results of Scenario 2 concerns the specification of a causal antecedent to the focal event in the gastric disturbances versions—namely the elimination of caffeine intake. The version stated “The doctor tells you that if you eliminate caffeine there is a 40% chance your gastric disturbances will stop.” Although the causal antecedent (for the stoppage of the disturbances) was presupposed by the doctor when making the forecast, perhaps positive thoughts associated with taking action to improve a condition influenced participants' perceptions of certainty. Identifying all ways in which context can influence perceptions of certainty is beyond the scope of this article, but it appears, now more than ever, that such an endeavor is an important one. Experiments 1 and 2, which demonstrate context effects with a novel forecast format and with two types of response-scale formats, suggest that context effects for judgments of certainty might be more pervasive than previously supposed.

Experiment 3a

Thus far, we have argued that the context manipulations used in Experiments 1 and 2 influenced perceptions of uncertainty because the manipulations varied the representativeness of the event to the context. Furthermore, we have argued that the effect of representativeness, because it is a product of associative processes, can operate independently of rule-based assessments of probability. Although the results of Experiments 1 and 2 are certainly consistent with these arguments, the base-rate account deserves further consideration.

We noted earlier that, for the scenarios used in Experiments 1 and 2, participants have no apparent basis for adjusting the experts' forecasts toward their own base-rate estimates. However, we cannot rule out the possibility that participants had some rationale for believing that their perceptions of the base rates should be used in conjunction with the experts' opinions (as is often the case in everyday life). We also cannot rule out a more implicit use of subjective base rates. Perhaps participants' perceptions of certainty were influenced by their subjective base rates in a manner consistent with a weighted averaging rule, even though the participants had no intention to use this information in such a way. With a base-rate interpretation still viable, the evidence for the proposal that associative processes can influence uncertainty independently of rule-based

assessments is less than fully compelling. Experiment 3a was designed to demonstrate that representativeness and associative processing play a central and sufficient role in the context effects of the type observed in the present experiments. Specifically, we designed an experiment in which the influence of subjective base rates could be ruled out as an explanation for observed context effects.

It seems reasonable to assume that for any event–context pair, both subjective base-rate estimates and perceptions of representativeness are functions of the number of times people have seen the event paired with the context (prior to the experiment session). This assumption might appear to render subjective base rates and perceived representativeness inseparable. However, although representativeness and subjective base rates might be correlated because they share common variance with a participant's exposure to event–context pairings, they are not redundant concepts and were considered separable for the present experiment. In fact, the ability to separate the two concepts is the key feature of this experiment. To distinguish the role of representativeness from the role of base rates, we created a situation in which the influences of representativeness would drive participants' certainty responses in one direction, whereas the use of subjective base rates would drive responses in a different direction.

As in Experiments 1 and 2, Experiment 3a involved measuring participants' perceptions of certainty for events in specified contexts. Unlike Experiments 1 and 2, four of the scenarios contained two separate probability estimates (e.g., 35% by one forecaster and 50% by another), and seven other scenarios contained a range of probability estimates (e.g., 60–70% by one forecaster). By presenting two forecasts (or a range estimate) in each scenario, we created a very strong test of the base-rate account. This situation implicitly encouraged participants to use their prior probabilities (i.e., subjective base rates); when there is vagueness or disagreement in experts' opinions, it would seem reasonable that people use their own perceptions of probability to resolve the vagueness or disagreement.

For the critical part of Experiment 3a, participants read 11 scenarios that each specified a possible event and forecasts for the event. For a given scenario, some of the participants, those in a context-present condition, read a version that described a specific context (along with information about the event and forecasts). On the basis of pilot testing, we knew that the event would be perceived as highly representative of that context. Other participants, those in the context-absent condition, read a version that contained the same information about the event and forecasts but left the context unspecified. For example, some participants read a scenario that indicated that a forecaster had predicted tourists had a 10–30% chance of being bitten by a snake while on a tour of an Amazon rain forest. Other participants read that a forecaster had predicted a 10–30% chance of being bitten by a snake while on a tour in a particular location, but the location was not revealed. All participants provided certainty estimates on a verbal certainty scale.

The critical question was whether certainty estimates would be *higher* in the context-present condition than in the

context-absent condition. Although it might seem that this type of context effect could plausibly be attributed to the influence of either representativeness or base rates, we designed an additional feature of this experiment to ensure that the direction of this effect could not be attributed to the influence of subjective base rates. Namely, the values that were presented as forecasts in this experiment exceeded the subjective base rates of a majority of participants in the context-present condition. This was achieved by pretesting, in a separate sample, subjective base rates for the event in the given contexts (i.e., the contexts specified in the context-present condition); we then used the pretesting data to guide our selection of the probabilities to be inserted as expert forecasts for Experiment 3a. Given that we ensured that participants' subjective base rates were, on average, below the experts' forecasts, use of base rates by participants in the context-present condition would push their probability assessment downward from the experts' forecasts. This would cause perceptions of certainty to be *lower* in the context-present condition than in the context-absent condition, where participants' responses would reflect their interpretation of the experts' forecasts without any influence of context.

We did expect that subjective base rates would have a downward influence on certainty estimates, but we also expected that perceptions of representativeness would influence participants' certainty in the opposite direction. Although these two processes would largely cancel each other out, we predicted that certainty estimates would be higher in the context-present condition than in the context-absent condition. This effect would suggest that the influence of representativeness outweighed the influence of processes involving subjective base rates.

Method

Participants. The participants were 168 undergraduate students at the University of Iowa who were enrolled in an introductory psychology course.

Materials and procedure. We constructed two versions of 11 scenarios for the experiment. Each scenario included a description of a target event that might occur, as well as numeric forecast information. In the context-present version of each scenario, the target event was highly representative of the context. In the context-absent version, the same event was described, but the context was left unspecified. The numeric forecast information was identical in the two versions of each scenario. (In Scenarios 2, 4, 6, 7, 9, 10, and 11, numeric forecasts were presented in the form of a range estimate; in Scenarios 1, 3, 5, and 8, separate estimates from two forecasters were presented.) Appendix B contains summary information about all 11 scenarios. Scenarios 6 and 9 are shown below, with the parenthetical information varying between the context-present and context-absent versions:

Scenario 6. The United States Overseas Tourism Company has published a list of estimates of risks involved in traveling to various overseas countries. One of the risk categories is for terrorist bombings. The publication specifies how likely it is that a terrorist bombing would occur within a country during a 2-week time period, which is the typical length of a tourist's stay. [For the country of Israel/For one country], the publication states that there is a 15–25% chance that a bombing

would occur during a typical tourist's visit. How likely do you think it is that a bombing would occur [in Israel/in this country] during a typical tourist's visit?

Scenario 9. *Traveler Magazine* recently published an article titled "Travel Tips for the Fearful Traveler." The article listed estimates of the chances of experiencing several types of frightful events at numerous tourist destinations around the world. For example, for a tourist taking a week-long tour [of an Amazon rain forest/in one location], the article suggested that his or her chance of being bitten by a snake was somewhere between 10 and 30%. Imagine that a person named Peter is about to take a week-long tour [of an Amazon rain forest/in that location]. How likely do you think it is that Peter would be bitten by a snake on his trip?

For each scenario, participants provided responses on the same type of scale that was used in Experiment 1, but the length of the scale line was 130 mm instead of 150 mm. Two different orders were used for the presentation of the scenarios, but the order variable did not interact with the context variable and is not discussed further.

Pretesting the Representativeness of Event–Context Pairs Used in Experiment 3a

Prior to conducting Experiment 3a, we tested the 11 event–context pairs used in the experiment to ensure that they were in fact high in perceived representativeness. In all, the pretesting included 44 event–context pairs—11 different events each paired with four unique contexts. We expected the event to be judged as highly representative of one context, unrepresentative of a second context, and moderately representative or unrepresentative of the other two contexts. For example, we pretested "snake bites–Amazon rain forest" (high representativeness), as well as "snake bites–Nebraska corn fields," "snake bites–Colorado Rockies," and "Snake bites–Swiss Alps." Ninety-two participants saw these pairs in one of two random orders. Participants were told to "look at each pair and quickly decide the degree to which you think the event is typical of, or generally associated with, the location." Responses were made on a 9-point scale anchored by -4 (*very atypical*), 0 (*neutral*), and $+4$ (*very typical*). The results of the pretesting confirmed that our high-representativeness pairs were, in fact, judged by participants as highly typical. The mean rating was $+3.17$ ($SD = 0.86$) for the high-representativeness pairs that were used in the scenarios of Experiment 3a. The mean rating for the pairs that were designed to be nonrepresentative was -0.56 ($SD = 1.00$).

Pretesting the Subjective Base Rates of Events Used in Experiment 3a

Also prior to conducting Experiment 3a, we assessed the subjective base rates of the events within the contexts that were used in the experiment. We used the pretesting results to determine the numbers that would be cited as forecasts in the scenarios of Experiment 3a. The participants in the pretesting read the context-present versions of the scenarios that were designed for Experiment 3a, but the forecast information was removed from those scenarios. Participants were asked to provide a numeric estimate of the probability that the target event would occur. In other words, participants provided their subjective base rate for the target event.

For example, to determine their subjective base rates for the events in Scenarios 6 and 9, participants read and responded to the following:

Scenario 6. The United States Overseas Tourism Company has published a list of estimates of risks involved in traveling

to various overseas countries. One of the risk categories is for terrorist bombings. The publication specifies how likely it is that a terrorist bombing would occur within a country during a 2-week time period, which is the typical length of a tourist's stay. How likely is it that a bombing would occur in Israel during a typical tourist's visit? There is a ____% likelihood.

Scenario 9. Traveler Magazine recently published an article titled "Travel Tips for the Fearful Traveler." The article listed estimates of the chances of experiencing several types of frightful events at numerous tourist destinations around the world. For example, the article provided an estimate for the chance that a tourist taking a week-long tour of an Amazon rain forest would be bitten by a snake. Imagine that a person named Peter is about to take a week-long tour of an Amazon rain forest. What is the likelihood that Peter would be bitten by a snake on his trip? There is a ____% likelihood.

Base-rate estimates for the scenarios were collected from 77 participants. For each scenario, we determined the base-rate estimate that fell at the 67th percentile. This estimate became the midpoint of the forecasters' estimates in Experiment 3a. For example, the pretest data indicated that the 67th percentile for the perceived base rate of a snake bite on a tour of an Amazon rain forest was 20%. Consequently, we inserted 10% and 30% as forecaster estimates in Experiment 3a. This use of pretesting data ensured that, on average, 67% of the participants in Experiment 3a read forecaster estimates that exceeded their own subjective base rate for a given event. In other words, for any given scenario, two thirds of the participants in Experiment 3a could be expected to have personal base-rate estimates that were lower than the average of the two expert forecasts.

Results

The critical question regarding these results is whether participants expressed more certainty in the context-present condition than in the context-absent condition. Participants' uncertainty responses for the 11 scenarios were scored in millimeters, from 0 to 130. These scores were submitted to a MANOVA with context (present vs. absent) as a between-subjects factor. As predicted, the overall analysis showed that participants in the context-present condition did in fact express more certainty ($M = 64.47$) than participants in the context-absent condition ($M = 62.14$), $F(11, 155) = 2.28$, $p = .01$. Table 3 displays means and standard deviations for the two versions of each scenario, as well as effect-size estimates for the differences between means. As can be seen from Table 3, the overall effect was driven primarily by the results of Scenarios 6, 9, and 11. Univariate comparisons revealed significant differences between context-present and context-absent versions for these three scenarios (all p s < .05) but not the others (all p s > .05).

Hence, there is clear evidence from three scenarios that the addition of context information boosted participants' certainty that the target event would occur. The direction of this effect cannot be explained by the base-rate account. The pretesting of the stimulus materials ensured that, on average, participants' subjective base rate for the event fell below the forecasted probabilities. Therefore, the influence of base rates worked *against* the detection of the observed effect. This suggests that the magnitude of the overall effect detected here might underestimate the influence of representativeness. This is also a possible consideration when

Table 3
Mean Uncertainty Responses and Effect-Size Estimates for Scenarios of Experiment 3a

Scenario	Context				<i>d</i>
	Present		Absent		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1	61.0	22.9	59.6	18.6	0.07
2	80.4	20.0	82.2	19.7	-0.09
3	96.1	16.4	100.6	18.7	-0.26
4	73.5	15.6	74.2	12.3	-0.05
5	72.1	16.7	72.5	16.7	-0.02
6	41.8	22.2	32.9	21.7	0.41
7	72.4	20.3	70.2	18.9	0.11
8	62.3	22.1	61.7	19.1	0.03
9	39.9	20.4	28.6	16.3	0.61
10	72.0	20.7	72.7	16.9	-0.04
11	37.7	23.2	28.1	17.9	0.46

Note. The uncertainty responses were scored in millimeters; the maximum score was 130. The *d* column displays the standardized mean difference in responses between the context-present and context-absent versions of the scenarios.

interpreting the variability in the size of the effects across scenarios. For scenarios showing no effect for the context manipulation, it is possible that the high representativeness of an event-context pair pushed certainty upward but that this influence was offset by the downward push of subjective base rates.

Before drawing further conclusions from this finding, we should rule out another alternative explanation. It might be appropriate to question whether the specification of any context, and not necessarily highly representative context, inflates people's perceptions of certainty. If this were the case, then the specification of only a moderately representative or unrepresentative context should produce effects that were similar to those seen in Experiment 3a. We conducted an additional experiment to test this possibility.

Experiment 3b

Experiment 3b was identical to Experiment 3a with two important exceptions. First, the contexts were changed such that the target events were not representative of the described contexts. Second, the forecasts in the scenarios were changed accordingly, ensuring that (just as in Experiment 3a) participants' subjective base rates for the events in the new contexts were again lower on average than the experts' forecasts.

As a result, context representativeness would not push participants' certainty in an upward direction. The influence of representativeness on associative processing was neutral or perhaps negative. If certainty estimates in the context-present condition were higher than those in the context-absent conditions, this result would seriously undermine our conclusion that the observed context effects are attributable to representativeness rather than the presence of some context per se.

Method

Participants. The participants were 85 undergraduate students at The Ohio State University who were enrolled in an introductory psychology course.

Materials and procedure. The procedures were identical to those of Experiment 3a. As mentioned above, the only differences in materials involved the contexts of the target events and the experts' numeric forecasts. We selected the new "nonrepresentative" contexts from the event-context pairs that were pretested for Experiment 3a. Recall that the mean typicality ratings for the event-context pairs used in Experiment 3a was +3.17 ($SD = 0.86$); the mean typicality ratings for the event-context pairs used in Experiment 3b was -0.44 ($SD = 1.00$). Based on these figures, it would be difficult to argue that the scenario events were representative of the contexts used in this experiment. The new forecasts that were cited in the scenarios were derived by pretesting the base-rate perceptions of a separate sample of 27 Ohio State University students. The procedures for this pretesting were identical to those used in the pretesting for Experiment 3a. As was the case for Experiment 3a, the means of the experts' forecasts were at the 67th percentile of the subjective base-rate estimates of the target events in our pretest sample. Summary information about the 11 scenarios is shown in Appendix C.

Results

The results of Experiment 3b were consistent with our conclusion from Experiment 3a and, in fact, bolstered that conclusion. Participants in the context-present condition expressed significantly less certainty ($M = 39.81$) than participants in the context-absent condition ($M = 50.85$), $F(11, 73) = 7.26, p < .001$. Table 4 displays the means and standard deviations for the two versions of each scenario, as well as effect-size estimates for the differences between means. These results clearly rule out the idea that the presence of context itself, not the influence of representativeness, can account for the results of Experiment 3a.

Table 4
Mean Uncertainty Responses and Effect-Size Estimates for Scenarios of Experiment 3b

Scenario	Context				<i>d</i>
	Present		Absent		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1	50.6	14.3	63.2	19.3	-0.07
2	62.7	9.4	62.7	22.0	0.00
3	69.4	8.6	76.3	19.0	-0.47
4	40.9	18.8	57.0	21.5	-0.80
5	40.5	14.1	54.3	23.4	-0.71
6	16.8	10.1	30.0	16.7	-0.96
7	30.2	20.7	37.8	24.5	-0.34
8	30.6	14.0	59.0	19.4	-1.68
9	19.2	14.8	23.8	14.1	-0.32
10	50.7	17.1	64.5	23.7	-0.67
11	26.4	13.3	30.9	22.3	-0.25

Note. The uncertainty responses were scored in millimeters; the maximum score was 130. The *d* column displays the standardized mean difference in responses between the context-present and context-absent versions of the scenarios.

Discussion of Experiments 3a and 3b

In Experiment 3a, the presence of context increased perceived certainty (relative to the context-absent condition), whereas in Experiment 3b, the presence of context decreased perceived certainty. The base-rate account cannot readily explain this pattern of findings. The scenarios in both experiments were constructed such that the midpoint of the forecasters' estimates fell above the subjective base rates for two thirds of the participants. Hence, if subjective base rates were solely responsible for the robust effects in Experiment 3b, then we would expect to observe the same effects in Experiment 3a. The representativeness account readily explains the difference in Experiments 3a and 3b by assuming that the low representativeness of the contexts in Experiment 3b contributed to the downward push on perceived certainty and/or the high representativeness of the contexts in Experiment 3a pushed certainty upwards, in opposition to any influence of subjective base rates.

Experiment 4

One potential criticism of Experiments 3a and 3b concerns the possible influence of base rates in the context-absent condition. In that condition, the context is not disclosed to the participants. For example, in the context-absent version of Scenario 9, participants were given a forecaster's estimates for getting a snake bite "in one location." Given the absence of context information, there would appear to be no relevant base rate influencing participants' perceptions of certainty. Nevertheless, one might argue that participants formulated a base-rate estimate for the very general, no-context events, and that this estimate was combined with the forecast information to determine participants' certainty judgments. For the snake-bite scenario, participants might have generated a general base rate for getting a snake bite in all possible vacation locations, which would probably be lower than the subjective base rates for snake bites in the location specified in the context-present condition of Experiment 3a (i.e., the Amazon rain forest). This could account for the observed pattern of results in Experiment 3a (and Experiment 3b if one assumes that the general base rates in the context-absent condition would be greater than the specific base rates in the context-present conditions).

Although this account is certainly possible, it seems implausible when considered from a participant's point of view. The scenarios make it clear to participants that the forecasters' estimates are specific to the particular context, and participants know that this context information is being withheld from them. It seems unlikely that participants considered a general base-rate estimate (e.g., the rate of snake bites in all possible vacation spots) relevant to interpreting the forecasts of experts who knew the exact context. It seems more plausible that participants ignored considerations of context and based their responses on the forecast information. Nevertheless, the base-rate account could be extended to suggest that general subjective base rates have a pervasive influence on certainty, even when a

research participant doesn't consciously consider those general base rates to be relevant to a certainty judgment. Any experiment using a context-absent condition as a comparison or control condition could be susceptible to this extended form of the base-rate account. Therefore, we designed an experiment that was similar to 3a and 3b but did not involve a context-absent condition.

Participants read a high- and low-representativeness version of 10 scenarios (similar to those used in Experiments 3a and 3b). Instructions informed participants that they would be reading event descriptions containing differing numeric forecast estimates from two experts and that their job was to decide—based on the forecast information and their own knowledge relevant to the described event—what would be the best single probability estimate for the event. Although we anticipated that changing to a numeric dependent measure would increase participants' use of the numeric forecasts and lessen their sensitivity to the context information, the task instructions essentially encouraged participants to use context information in determining their judgment. Hence, we expected to see strong context effects. The pertinent question was whether participants would use context information in a manner consistent with the base-rate account or the representativeness account.

We used pilot testing to determine the median subjective base rates separately for the high- and low-representativeness versions of each scenario. These median estimates were then used to determine what forecast information the participants in Experiment 4 would see. For example, if the median estimate for a low-representativeness version was 35% in pilot testing, the participants in the main experiment saw forecasts such as 30% and 40% embedded in that low-representativeness version. This ensured that when the participants in the main experiment read the scenario, half would have subjective base-rate estimates that fell above the midpoint of the experts' forecasts and half would have subjective base-rate estimates that fell below that midpoint. Therefore, if participants used only their subjective base rates and the presented forecasts to help determine what the single best probability estimate was, then equal numbers of responses (i.e., probability estimates) should fall above and below the midpoint of the experts' forecasts. The base-rate account predicts that the proportions of people giving estimates above versus below the midpoint of the experts' forecasts would be roughly equivalent for the high- and low-representativeness versions of the scenarios.

If, however, the representativeness of an event–context pair had an influence that was partially independent of subjective base rates, a different pattern of results would be expected. Specifically, the proportion of people giving estimates above versus below the forecasts' midpoint would be significantly higher for the high-representativeness versions of the scenarios than for the low-representativeness versions of the scenarios.

Method

Participants. The participants were 60 undergraduate students at the University of Iowa who were enrolled in an introductory psychology course.

Design, materials, and procedure. The design of the experiment was completely within-subjects; each participant responded to a high- and low-representativeness version of 10 scenarios. We borrowed the basic story lines of these scenarios and the high- and low-representativeness contexts from the scenarios in Experiments 3a and 3b. Appendix D contains summary information about all 10 scenarios. Initial instructions informed participants that for each scenario, two experts had given differing probability estimates for how likely a described event was. The instructions also stated, "Your job is to decide what you think is the best single estimate of how likely it is that the event will happen." After reading a short description of the response scale being used, participants read the scenarios and provided a certainty response between 0% and 100% for each. The high- and low-representativeness versions of Scenario 3 are shown below.⁸

High representativeness for Scenario 3. A publication from the United States Tourism Company states that, for the country of Israel, there is a 20% chance that a terrorist bombing would occur somewhere in that country during any 3-day time period. The estimate given by another tourism company is 28%. Robert is traveling to Israel for 3 days to conduct business. Given what you know about terrorist bombings in Israel, and given the two likelihood estimates described above, what do you think is the best single estimate for the likelihood that there will be a terrorist bombing somewhere in Israel during Robert's visit? ____%

Low representativeness for Scenario 3. A publication from the United States Tourism Company states that, for the city of Toronto, Canada, there is a 1% chance that a terrorist bombing would occur somewhere in that city during any 2-week time period. The estimate given by another tourism company is 4%. Robert is traveling to Toronto for 2 weeks to conduct business. Given what you know about terrorist bombings in Toronto, and given the two likelihood estimates described above, what do you think is the best single estimate for the likelihood that there will be a terrorist bombing somewhere in Toronto during Robert's visit? ____%

The order in which participants saw the scenarios was arranged such that a given participant first responded to one version of each of the 10 scenarios (some high- and some low-representativeness versions) and then responded to the second version of each scenario. For any given scenario, half the participants saw the low-representativeness version first, and half saw the high-representativeness version first.

The expert forecasts included in the scenarios were based on pretesting data. Specifically, the midpoint of the two forecasts was the median subjective base rate reported by pretesting participants.

Pretesting of Subjective Base Rates for Experiment 4

In the pretesting, we solicited (from 40 students at the University of Iowa) subjective base-rate estimates for the event–context pairs that would be used in Experiment 4. Pretesting participants saw the exact same event and context information as participants in Experiment 4, with the exception that the numeric forecasts were

⁸ In Scenario 9, the time period specified in the high-representativeness version (3 days) was shorter than the time period specified in the low-representativeness version (2 weeks). This was done to help avoid floor effects for the low-representativeness versions and ceiling effects for the high-representativeness versions. Similar precautions were taken for five other scenarios (see Appendix D). For all scenarios, the time periods that were used in the main experiment were identical to those that were pilot tested.

not revealed. For example, the high- and low-representativeness versions of Scenario 3 are shown below.

High representativeness for Scenario 3. A publication from the United States Tourism Company provided an estimate for the chance that a terrorist bombing would occur somewhere in Israel during any 3-day time period. A second estimate was given by another tourism company. Robert is traveling to Israel for 3 days to conduct business. Given what you know about terrorist bombings in Israel, what do you think is the best single estimate for the likelihood that there will be a terrorist bombing somewhere in Israel during Robert's visit? _____%

Low representativeness for Scenario 3. A publication from the United States Tourism Company provided an estimate for the chance that a terrorist bombing would occur somewhere in Toronto, Canada, during any 2-week time period. A second estimate was given by another tourism company. Robert is traveling to Toronto for 2 weeks to conduct business. Given what you know about terrorist bombings in Toronto, what do you think is the best single estimate for the likelihood that there will be a terrorist bombing somewhere in Toronto during Robert's visit? _____%

As was the case for participants in Experiment 4, the pretesting participants were informed, "Your job is to decide what you think is the best single estimate of how likely it is that the event will happen." The pretesting participants were also given the same instructions about the response scale being used; responses were made on a 0–100% scale. Finally, the same design and counterbalancing schemes that were used in Experiment 4 were also used in the pretesting. As stated above, we used the median response for each scenario version to determine the midpoint of the forecast estimates used in Experiment 4. This ensured that for any given scenario version in Experiment 4, the number of participants having subjective base rates that fell above versus below the midpoint of experts' forecasts would be equivalent.

Results and Discussion

We classified each response for each scenario version as below, at, or above the midpoint of the forecasters' estimates. Summary data are presented in Table 5. The critical question regarding the results is whether the proportion of people giving estimates above versus below the midpoint of the experts' forecasts was significantly higher for the high-representativeness versions of the scenarios than for the low-representativeness versions of the scenarios. Across all participants, 445 responses in the low-representativeness versions fell below the relevant midpoints and 100 responses fell above. In contrast, only 252 of the responses in the high-representativeness versions fell below the relevant midpoints, while 233 fell above. A chi-square analysis of these overall data indicate that the proportion of responses falling above versus below the midpoints was significantly different in the predicted direction for high- and low-representativeness versions, $\chi^2(1, N = 1030) = 103.4, p < .0001$.⁹ Analogous chi-square analyses for individual scenarios were significant for all 10 scenarios in the predicted direction.

One unanticipated aspect of these data was that for the high-representativeness versions, participants were about equally likely to provide a response that was above versus below the midpoint of the forecasters' estimates, whereas for the low-representativeness versions, participants showed a clear tendency to provide responses that were below the

Table 5
The Number of Responses Falling Below, At, and Above the Midpoint of the Forecasters' Estimates for Each Version of Each Scenario in Experiment 4

Scenario	Version					
	Low representativeness			High representativeness		
	Below	At	Above	Below	At	Above
1	45	5	9	32	1	27
2	44	7	9	15	18	26
3	55	1	4	24	6	29
4	34	18	7	23	21	16
5	42	0	17	9	27	24
6	34	1	25	11	22	27
7	50	0	10	25	18	17
8	46	12	2	38	0	22
9	44	0	15	33	0	27
10	51	7	2	42	0	18
Sum	445	51	100	252	113	233

Note. There were 60 responses (sometimes 59 because of missing data) for each version of each scenario. Chi-square analyses for each scenario indicate that the proportions of responses falling below (vs. above) the experts' midpoint were significantly greater for the low-representativeness versions than for the high-representativeness versions.

forecasters' midpoint. At first glance, this asymmetry in the results seems to suggest that perceptions of high representativeness have no influence on certainty, whereas perceptions of low representativeness have a very powerful influence on certainty. This conclusion, however, is not warranted. Any factor that would produce an overall lowering of participants' estimates would produce this type of asymmetry (e.g., an implicit or explicit assumption by some participants that forecasters tend to overpredict events, or a simple bias toward the lower of the two expert forecasts). A factor that produces a main-effect lowering of participants' responses would increase the proportions of responses falling below the experts' midpoints for both the low- and high-representativeness versions—making it appear that low-representativeness contexts had more influence on responses than did high-representativeness contexts. Exploring this asymmetry in the results for low- and high-representativeness contexts warrants future work, but none of the possible explanations for this asymmetry conflicts with the main conclusion from this Experiment 4: Differences in perceived representativeness can produce differences in certainty that cannot be explained by base-rate accounts.

⁹ For simplicity, the chi-square tests reported here excluded the responses that fell exactly at the midpoint of the experts' forecasts. However, these could be considered meaningful responses that should be considered in the analyses. Alternative analyses (not reported in the text) that included the midpoint support the same conclusions as the described chi-square analyses. For one analysis, responses below, at, and above the midpoint were scored as $-1, 0$, and 1 , respectively. We then used these scores to compute total scores for the low-representativeness versions and for the high-representativeness versions for each participant. A paired t test revealed that these total scores were significantly different.

A reviewer of this article noted that the observed differences in responses for the high- and low-representativeness versions of the scenarios in Experiment 4 were confounded with the size of the forecast midpoints. The forecasts in the high representativeness versions were generally higher than those in the low-representativeness versions. This confound is a natural one—highly representative events will tend to elicit higher forecasts than unrepresentative events. Nevertheless, the confound leaves open the possibility that the results of Experiment 4 could be explained by assuming that when the forecasts were low, participants tended to respond lower, and when the forecasts were high, participants tended to respond higher. Although this potential alternative explanation cannot apply to Experiments 1, 2, 3a, and 3b, we conducted an additional study to rule out this explanation for Experiment 4.

In this additional study, a replication of Experiment 4, we removed the overall differences in the size of the forecasters' estimates between the high- and low-representativeness versions of the scenarios. This was accomplished by altering aspects of the scenarios (usually the stated time frame for when the focal event might occur) to draw the pretested base rates in the low-representativeness versions up and push the base rates in the high-representativeness versions down. For example, we altered the low-representativeness version of Scenario 3 to discuss the chances of a terrorist bombing in Toronto during a 1-year time frame, while the high-representativeness version discussed the chances of a terrorist bombing in Israel in a 3-day time period. The pretest data for the replication revealed that the pretested base rates for the high- and low-representativeness versions were roughly equivalent for 8 of the 10 scenarios.¹⁰ As in Experiment 4, we used the medians of the pretested base rates to determine the midpoints of the forecasters' estimates in the replication.¹¹ Hence, across the 8 scenarios, there was no longer a confound between the size of the forecasters' estimates and the representativeness of the scenario version (the mean forecasts included in the high and low versions were 22.0% and 21.6%, respectively). Nevertheless, participants showed the same robust, albeit smaller, difference between how they responded (below or above the forecasts' midpoint) to the high- versus low-representativeness versions, $\chi^2(1, N = 390) = 7.56, p < .01$.¹² The replication results also exhibited the same asymmetry as was noted for Experiment 4. If we exclude responses that fell exactly at the midpoint of the forecasters' estimates, 79.8% of the remaining responses for the low-representativeness versions fell below the midpoint, whereas 67.5% of the remaining responses for the high-representativeness versions fell below the midpoint. Again, this asymmetry deserves attention in future work.

The results of Experiment 4 and its replication provide clear support for the representativeness account; the extended form of the base-rate account cannot account for the observed effects. When deciding what the best single probability estimate for a given event was, participants appeared to make strong use of the forecasts given by the experts; across all scenarios in Experiment 4, 71% of responses fell within the range defined by the forecasters' estimates listed in a scenario. Participants were also encour-

aged, however, to use their own relevant knowledge in deriving the best estimate. In making adjustments from these forecaster estimates, it is possible that participants used their subjective base rates. These adjustments, however, cannot account for the context effects that were observed, because the median subjective base rate fell at the midpoint of the experts' forecasts for both the high- and low-representativeness contexts. Consistent with the representativeness account, it appears that the context effects were observed because people's judgments of certainty were influenced by the representatives of an event for a given context, independently of the influence of subjective base rates.

General Discussion

Explaining Context Effects

In combination, the present experiments provide strong support for our representativeness account of how contextual information can influence perceptions of certainty. Experiments 1 and 2 demonstrated that context effects are not restricted to situations in which forecast information is vague—an assumption made by the original base-rate account of context effects. Perceptions of certainty were shown to be sensitive to context manipulations even when precise numeric forecast information was known. Experiments 3a, 3b, and 4 demonstrated that the base-rate account cannot explain all forms of context effects. More specifically, whether the base-rate account of context effects is interpreted as a process account (intended to describe the critical processes yielding the effects) or as a functional account (intended to model the observed pattern of results), it cannot explain the pattern of results observed in Experiments 3a, 3b, and 4.

Our research does not address the validity of explanations that have been offered for previous findings involving verbal forecasts (e.g., Wallsten et al., 1986; Weber, 1994; Weber & Hilton, 1990). However, the account we have proposed provides a substantially different explanation for context effects. Whereas earlier accounts described context effects as consequences of a weighted averaging of vague probability estimates (Wallsten et al., 1986) or an adaptive consideration of outcome severity (Weber, 1994; Weber & Hilton, 1990), the representative account assumes that context effects can be products of relatively simple event-context associations that can influence certainty independently of more rule-based considerations. It is quite plausible that

¹⁰ The pretested base rates for Scenarios 2 and 5 remained extremely disparate between high- and low-representativeness versions despite alterations to the time frames. However, conclusions from the overall analysis of the Experiment 4 replication are unaffected by whether data from these scenarios are included or excluded.

¹¹ In the replication, we also ensured that the size of the difference between the two forecasters' estimates in each scenario was equivalent in the high- and low-representativeness versions.

¹² We also conducted a paired *t* test, analogous to the one described in Footnote 9. It was significant and supports the same conclusion as the chi-square analysis.

such associative processes played important roles in the context effects previously demonstrated for verbal forecasts; Experiment 4 demonstrated that even when numeric probability estimates were solicited from participants, as was done in previous context-effects studies, the resolution of the discrepant forecast information was determined, at least in part, by perceptions of representativeness.

One important implication of this new view of context effects is that when people attempt to deliver precise information about the probability of an event, the communicated information can have more than its intended meaning. For example, imagine a doctor who informs her patient that he has a 70% chance of a full recovery from a knee surgery. Although the patient may accept that numeric probability as an appropriate forecast, the doctor might have also communicated information that could affect the patient's more associatively based thoughts and feelings about the possibility of recovery. If the doctor mentioned positive reasons for why there is a 70% chance of a full recovery, the patient might have a greater feeling of optimism about the surgery than if the doctor mentioned negative reasons for the 70% estimate (see Teigen & Brun, 1999, for a related point). Communicators should not assume that a probability forecast that is perceived as precise and credible is the only determinant of a receiver's certainty.

Representativeness and Base Rates

The present findings also suggest an important addition to our understanding of the construct of representativeness. Representativeness has been assumed to be a judgment based on similarity, and such judgments have been assumed to be powerful mediators of people's estimates of probability. In fact, classic conceptions of representativeness view the judgment as a tool that people can purposefully use to make probability estimates (Kahneman & Tversky, 1972; Tversky & Kahneman, 1974, 1982), and many studies on base-rate neglect and conjunction-rule violations were intended to demonstrate that people overuse representativeness to the detriment of other factors that should be considered in estimating probability (e.g., Kahneman & Tversky, 1973; Tversky & Kahneman, 1983).

We do not disagree with the idea that representativeness can serve as a tool for generating certainty estimates. However, our analysis suggests that representativeness can influence the perceptions of certainty in ways that are partially independent of a person's beliefs in objective probability. A person who believes there is a 5% chance of rain can feel more concerned when in London than in Madrid, because rain is more representative of London than Madrid. This person did not purposefully use representativeness as a tool but nevertheless was affected by the rain-London association (see Lopes & Oden, 1991, for a related argument). In the pretesting of base rates for Experiments 3a, 3b, and 4, the pretesting participants may or may not have used representativeness in the classic sense—as a purposefully used tool for estimating their subjective base-rate probability. However, the influence of representativeness in producing the context effects like those in

Experiment 1 was importantly different. Perceptions of representativeness were not purposefully used tools but rather were nondeliberative influences on perceived certainty.

This perspective raises some intriguing new questions about classic representativeness problems. For example, imagine a lawyer-engineer problem in which the jar is described as containing descriptions of 20 lawyers and 80 engineers (Kahneman & Tversky, 1973). Also imagine that a savvy respondent reads a lawyer-like description, appropriately attends to the base-rate information, and indicates there is a 50% chance that the described person is a lawyer and a 50% chance that the person is an engineer. Assuming some independence between beliefs in objective probability and a more implicit influence of representativeness, this respondent, despite the well-reasoned 50–50 response, might have an “intuitive hunch” that the person is probably a lawyer, not an engineer.

The present findings are also relevant to the issue of base-rate neglect. The observed context effects can be viewed as evidence that people's perceptions of certainty are sensitive to the *objective* base rate for an event in a given context; in general, certainty estimates were higher for events with high versus low *objective* base rates. Experiments 3a, 3b, and 4 indicate that this sensitivity to objective base rates was more likely mediated by the strength of event-context associations than by processes in which subjective base rates are averaged with forecast information. In other words, although responses appear to show sensitivity to objective base rates, this does not mean that participants made any attempt to estimate the objective base rates for the events in the contexts. The strengths of associations between an event and context—which would tend to be strong for events with high base-rates (snow in Colorado) and weak for events with low base rates (snow in South Carolina)—can account for the sensitivity to objective base rates.

This suggests that there may be instances for which the use of base-rate information is a natural property of associative processing, if the events in question are learned in association with a given context. Perhaps people have trouble using base rates in problems like the lawyer-engineer problem because the only way to use base-rate information is to purposefully apply abstract formal rules; the presentation of base-rate information in summary statistics (e.g., 20 engineers and 80 lawyers) functionally precludes any associative processing of base-rate information. Weber, Böckenholt, Hilton, and Wallace (1993) showed that doctors have no problem incorporating base-rate information into their diagnostic judgments when they can do so in a fashion that uses symptom-diagnosis associations from their memory base, even though they fail to appreciate the significance of base rates when presented with problems that require the rule-based integration of numeric base-rate information (Eddy, 1982).

Associative and Rule-Based Processing

The findings of the present research are consistent with Sloman's (1996) broad distinction between associative and

rule-based processing. As described earlier, a key component of Sloman's proposal is that the two processing systems are semi-independent. Relatedly, the representativeness account that was tested and supported here assumes that the associative processes that underlie the perceptions of representativeness can operate independently of a person's beliefs in the objective probability of an event and can be at least partially separated from the influence of subjective base rates. Also, if we adopt Sloman's proposal that rule-based processes involve the execution of rules that a respondent assumes is relevant to the task, then the context effects observed here do not appear to be the result of rule-based processing. The most obvious rule for modifying a forecaster's estimate would be to adjust in the direction of one's own subjective base rate, but this type of process cannot account for the observed findings.

There are several components of Sloman's proposal that were not tested here (e.g., automaticity, flexibility, sources of knowledge, nature of representation). Hence, the present research cannot be considered a broad test of Sloman's proposal. However, in line with Sloman's distinction, the present findings suggest that there may be important benefits to hypothesizing and exploring associative processes that might have effects that are independent of the more apparent rule-based processes. Furthermore, we suggest that an improved understanding of judgment and decision processes will require closer attention to associative processes, and this requires closer attention to the basic memory processes that underlie initial judgments and affect ultimate decisions (see Dougherty, Gettys, & Ogden, 1999; Weber et al., 1993; Weber, Goldstein, & Barlas, 1995; Weber, Goldstein, & Busemeyer, 1991).

In the context of the judgments described in this article, a memory-based approach would not explain why rain seems more likely in London than Madrid (given a 5% estimate) by focusing on the role of subjective base rates. Rather, a memory-based approach might assume that London and rain are closely associated in memory (e.g., rain is part of a London prototype), and activation of the concept of London enhances the activation of the concept of rain. A strong activation of "rain" might not directly conflict with a belief that there is a 5% chance of rain, because there is no direct mapping between activation strength (or construct accessibility) and the mental representation of a numeric probability estimate. In other words, it is possible that a person's judgment of certainty can be influenced by two mental representations that are quite distinct—one based on activation strength and tied to associations in memory, and another based on a belief in the objective probability and tied to knowledge that is deemed by the person to be relevant to determining the objective probability.

How would the influences of these two representations be combined? This question cannot be precisely answered with the available data, but this research does suggest an important constraint on possible answers. The present findings suggest that the combination of these two representations could not be explained by a simple weighted averaging model. In other words, the findings do not support the idea that perceptions of certainty based only on representative-

ness and perceptions of certainty based only on forecasters' estimates are averaged to form overall certainty estimates. This hypothesis would be indistinguishable from the base-rate account that was ruled out.

We speculate here about one process explanation for how representativeness would exert its influence on a response made under the conditions of the present experiments. On the basis of considerations of the forecasters' estimates (possibly including assumptions about biases of forecasters and assumptions about subjective base rates), a respondent develops a vague sense of a plausible range for the objective likelihood of the specified event. The representation of this plausible range would presumably depend on the nature of the response being solicited. For example, the respondent would establish a vague numeric range if a numeric probability is requested. Independently, and based on associative processes, the respondent would also experience, in a more passive sense, perceptions of whether the event is representative or unrepresentative for the context. Rather than quantifying these perceptions and then averaging them with the range of numbers that seemed plausible for objective likelihood, the respondent would be influenced by perceptions of representativeness in a more direct way. When the event seems very representative, the respondent would be biased toward the upper portions of his or her plausible range, and when the event seems very unrepresentative, the respondent would be biased toward the lower portions of his or her plausible range. This would mean that, even in a situation in which a respondent's plausibility range is already high, a highly representative context can boost it higher, and even in a situation in which a respondent's plausibility range is already low, a very unrepresentative context can deflate it further.

Determining the validity of this specific process account requires further research. More generally, determining how disparate thoughts about certainty—for example, one derived from associative-based thinking and one derived from rule-based thinking—combine to influence judgments and behaviors warrants further investigation. Consider a not uncommon example of a person who feels uneasy about the possibility of a plane crash, yet at the same time, knows that the objective likelihood of a crash is extremely small. Researchers know relatively little about whether and how such perceptions are combined to influence overall certainty and relevant behaviors.

Studying People's Judgments and Decisions Using Content-Impoverished Stimuli

A foundational assumption for much research on judgment and decision making is that real-life decisions can be represented as gambling decisions, and that researchers can study people's preferences among simple gambles using content-impoverted stimuli in order to understand the processes that guide people's decisions in real-world settings (Savage, 1954; von Neumann & Morgenstern, 1947). There is an ongoing discussion about the degree to which this gambling metaphor is useful and appropriate (see, e.g., Beach & Mitchell, 1987; Erev, Bornstein, & Wallsten, 1993;

Goldstein & Weber, 1995; Hastie, 1991; Rettinger & Hastie, 1998). The present demonstrations provide additional fodder for those theorists who wish to limit the use of the gambling metaphor. The fact that context influences uncertainty even when event probabilities are numerically specified suggests that the influences of stimulus content on real-world decisions cannot be fully understood by observing the effects of adjustments to the content-impoverished probabilities of gambles.

More important, however, the associative versus rule-based distinction discussed here might provide a helpful framework for understanding when the gambling metaphor is appropriate and useful, and when it is problematic. The great benefit of the metaphor is that it allows researchers to posit general theories of decision making—theories that are not tied to particular content domains. We think the metaphor is particularly useful when the behaviors of interest to a researcher are mediated by primarily rule-based processes. Rule-based processing treats information in a relatively abstract form and operates on it according to abstract and formal rules; content can often be ignored in rule-based processing. However, when the behavior of interest to a researcher is mediated by primarily associative processing, the gambling metaphor is problematic. With shifts in content come shifts in pre-experiment associations. By using content-impoverished stimuli, researchers remove the influence of associations and then fail to recognize that, although the associations cannot be fully represented in probability and outcome values, they do nevertheless drive uncertainty and behavior. When associations are measured or when content is manipulated, researchers can achieve a fuller perspective on the factors that influence judgments, decisions, and behaviors.

References

- Beach, L. R., & Mitchell, T. R. (1987). Image theory: Principles, plans, and goals in decision making. *Acta Psychologica, 66*, 201–220.
- Brun, W., & Teigen, K. H. (1988). Verbal probabilities: Ambiguous, context-dependent, or both? *Organizational Behavior and Human Decision Processes, 41*, 390–404.
- Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.
- Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual process theories in social psychology*. New York: Guilford Publications.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Denes-Raj, V., Epstein, S., & Cole, J. (1995). The generality of the ratio-bias phenomenon. *Personality and Social Psychology Bulletin, 21*, 1083–1092.
- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). *MINVERA-DM: A process model for judgments of likelihood*. *Psychological Bulletin, 106*, 180–209.
- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Epstein, S., Lipson, A., Holstein, C., & Huh, E. (1992). Irrational reactions to negative outcomes: Evidence for two conceptual systems. *Journal of Personality and Social Psychology, 62*, 328–339.
- Erev, I., Bornstein, G., & Wallsten, T. S. (1993). The negative effect of probability assessments on decision quality. *Organizational Behavior and Human Decision Processes, 55*, 78–94.
- Goldstein, W. M., & Weber, E. U. (1995). Content and discontent: Indications and implications of domain specificity in preferential decision making. In J. Busemeyer, R. Hastie, D. Medin (Vol. Eds.), *The psychology of learning and motivation: Vol. 32. Decision making from a cognitive perspective* (pp. 275–318). Academic Press.
- Hastie, R. (1991). A review from a high place: The field of judgment and decision making as revealed in its current textbooks. *Psychological Science, 2*, 135–138.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430–454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251.
- 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.
- Langer, E. J. (1989). *Mindfulness*. Reading, MA: Addison-Wesley.
- Lopes, L. L., & Oden, G. C. (1991). The rationality of intelligence. In E. Eells & T. Maruszewski (Eds.), *Rationality and reasoning*. Amsterdam: Rodopi.
- Miller, D. T., Turnbull, W., & McFarland, C. (1989). When a coincidence is suspicious: The role of mental simulation. *Journal of Personality and Social Psychology, 57*, 581–589.
- Rettinger, D. A., & Hastie, R. (1998, November). *Content and framing effects on decision making*. Paper presented at the Annual Meeting of the Society for Judgment and Decision Making, Dallas, TX.
- Russell, C. J., & Bobko, P. (1992). Moderated regression analysis and Likert scales: Too course for comfort. *Journal of Applied Psychology, 77*, 336–342.
- Savage, L. J. (1954). *The foundation of statistics*. New York: Wiley.
- Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1–66.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22.
- Teigen, K. H., & Brun, W. (1999). *The directionality of verbal probability expressions: Effects on decision, prediction, and probabilistic reasoning*. Manuscript submitted for publication.
- Tversky, A., & Kahneman, D. (1974, September). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1130.
- Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*, 293–315.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- Wallsten, T. S., Fillenbaum, S., & Cox, J. A. (1986). Base rate effects on the interpretations of probability and frequency expressions. *Journal of Memory and Language, 25*, 571–587.
- Weber, E. U. (1994). From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evalua-

- tion of uncertain outcomes and events. *Psychological Bulletin*, 115, 228–242.
- Weber, E. U., Böckenholt, U., Hilton, D. J., & Wallace, B. (1993). Determinants of diagnostic hypothesis generation: Effect of information, base rates, and experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1151–1164.
- Weber, E. U., Goldstein, W. M., & Barlas, S. (1995). And let us not forget memory: The role of memory processes and techniques in the study of judgment and choice. In J. R. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Decision making from a cognitive perspective: The psychology of learning and motivation* (Vol. 32, pp. 33–82). New York: Academic Press.
- Weber, E. U., Goldstein, W. M., & Busemeyer, J. R. (1991). Beyond strategies: Implications of memory representation and memory processes for models of judgment and decision making. In S. Lewandowsky & W. E. Hockley (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 75–100). Hillsdale, NJ: Erlbaum.
- Weber, E. U., & Hilton, D. J. (1990). Contextual effects in the interpretations of probability words: Perceived base rate and severity events. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 781–789.
- 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.

Appendix A

Summary Information for the 7 Scenarios Used in Experiment 1

The last sentences of each of the scenarios used in Experiment 1 are printed below. The information within brackets is the context information that was manipulated between versions of the scenarios (Versions A and B, respectively). Following each scenario excerpt is the numeric forecast information that was given in the scenario. This forecast information was described as being specific to the context that was described.

Scenario 1

Please mark on the rating scale below how likely you think it is that, by the year 2000, [beach volleyball/swimming] will be among the top three most popular sports for American adults. (50%)

Scenario 2

Please mark on the rating scale below how likely you think it is that there will be sufficient snow for the [Central Ohio/Central Colorado] ski race the week before Christmas. (35%)

Scenario 3

Please mark on the rating scale below how likely you think it is that Carol would pass Calculus 3, given that she has received [Cs/Bs] in the two previous classes. (70%)

Scenario 4

Please mark on the rating scale below how likely you think it is that David's [late-rising/early-rising] father will be in bed by 10 PM. (45%)

Scenario 5

Please mark on the rating scale below how likely you think it is that Tanya can transport her clients through [Des Moines/Los Angeles] during a rush hour without being slowed by more than 5 minutes. (20%)

Scenario 6

Please mark on the rating scale below how likely you think it is that Mary Beth will hand in her term paper [two days early/on time]. (60%)

Scenario 7

Please mark on the rating scale below how likely you think it is that Janet will contract malaria while in [Hawaii/India]. (30%)

Appendix B

Summary Information for the 11 Scenarios Used in Experiment 3a

The last sentences of each of the scenarios used in Experiment 3a are printed below. The information within brackets is the information that was manipulated between versions of the scenarios

(context-present and context-absent versions, respectively). Following each scenario excerpt is the numeric forecast information that was given in the scenario.

Scenario 1

How likely do you think it is that a man would be mugged [if he rode the New York City Subway for a week/in that hypothetical situation]? (35% and 50%)

Scenario 2

How likely do you think it is that [Topeka, Kansas/this location] will suffer damage from a tornado in the year 1998? (62–72%)

Scenario 3

How likely do you think it is that it will snow [in Moorhead, Minnesota, on the day of the event/on the day of the event]? (75% and 85%)

Scenario 4

The science categories are physics, chemistry, and biology. How likely do you think it is that [Harvard/the school] will win the competition? (45–60%)

Scenario 5

How likely do you think it is that [Los Angeles/the city] will experience an earthquake strong enough to be life threatening in the year 1998? (45% and 65%)

Scenario 6

How likely do you think it is that a bombing would occur [in Israel/in this country] during a typical tourist's visit? (15–25%)

Scenario 7

How likely do you think it is that Barbara will contract a form of malaria [while on her trip to Calcutta/while on her trip]? (50–60%)

Scenario 8

How likely do you think it is that [a drug addict who was arrested for dealing drugs/a person exhibiting this pattern] would commit murder in the 5-year time period after being released? (35% and 45%)

Scenario 9

Imagine that a person named Peter is about to take a week-long tour [of an Amazon rain forest/in that location]. How likely do you think it is that Peter would be bitten by a snake on his trip? (10–30%)

Scenario 10

How likely do you think it is that more than 20 people will die from heat stroke [in Arizona/in that state] in the summer of 1998? (55–65%)

Scenario 11

How likely do you think it is that within a 1-month time span—say the month of June, 1998—that a homicide will be committed in one of [the many United States Postal Buildings/the company's buildings]? (10–15%)

Appendix C

Summary Information for the 11 Scenarios Used in Experiment 3b

The last sentences of each of the scenarios used in Experiment 3b are printed below. The information within brackets is the information that was manipulated between versions of the scenarios (context-present and context-absent versions, respectively). Following each scenario excerpt is the numeric forecast information that was given in the scenario.

Scenario 1

How likely do you think it is that a man would be mugged [if he rode the Dubuque Metro Bus System for a week/in that hypothetical situation]? (25% and 40%)

Scenario 2

How likely do you think it is that [Peoria, Illinois/this location] will suffer damage from a tornado in the year 1998? (29–39%)

Scenario 3

How likely do you think it is that it will snow [in Leavenworth, Kansas, on the day of the event/on the day of the event]? (45% and 55%)

Scenario 4

The science categories are physics, chemistry, and biology. How likely do you think it is that [West Texas State/the school] will win the competition? (10–25%)

Scenario 5

How likely do you think it is that [Austin/the city] will experience an earthquake strong enough to be life threatening in the year 1998? (20% and 30%)

Scenario 6

How likely do you think it is that a bombing would occur [in Toronto, Canada/in this country] during a typical tourist's visit? (1–10%)

Scenario 7

How likely do you think it is that Barbara will contract a form of malaria [while on her trip to Washington/while on her trip]? (5–15%)

(Appendix C continues)

Scenario 8

How likely do you think it is that [a person with 100 or more overdue parking tickets/a person exhibiting this pattern] would commit murder in the 5-year time period after being released? (15% and 25%)

Scenario 9

Imagine that a person named Peter is about to take a week-long tour [of Nebraska/in that location]. How likely do you think it is that Peter would be bitten by a snake on his trip? (1–10%)

Scenario 10

How likely do you think it is that more than 20 people will die from heat stroke [in Indiana/in that state] in the summer of 1998? (25–35%)

Scenario 11

How likely do you think it is that within a 1-month time span—say the month of June, 1998—that a homicide will be committed in one of [the many national chain book stores in the United States/the company's buildings]? (3–8%)

Appendix D

Summary Information for the 10 Scenarios Used in Experiment 4 and Its Pretesting

The end of the last sentence for each low- and high-representativeness scenario used in Experiment 4 and its pretesting is printed below. (See the example scenario in the *Method* section of Experiment 4 to see what type of information preceded these excerpts.) Following each excerpt is the numeric forecast information that was given in Experiment 4 (not the pretesting) for that scenario.

Scenario 1, Low Representativeness

... what do you think is the best single estimate for the likelihood that a man would be mugged if he rode the Dubuque Metro Bus System once each night for a month? (2% and 10%)

Scenario 1, High Representativeness

... what do you think is the best single estimate for the likelihood that a man would be mugged if he rode the New York City Subway once each night for a week? (30% and 45%)

Scenario 2, Low

... what do you think is the best single estimate for the likelihood that Boulder, Colorado, will suffer damage from a tornado in 1999? (4% and 12%)

Scenario 2, High

... what do you think is the best single estimate for the likelihood that Topeka, Kansas, will suffer damage from a tornado sometime in the months of March, April, or May of next year? (60% and 80%)

Scenario 3, Low

... what do you think is the best single estimate for the likelihood that there will be a terrorist bombing somewhere in Toronto during Robert's visit? (1% and 4%)

Scenario 3, High

... what do you think is the best single estimate for the likelihood that there will be a terrorist bombing somewhere in Israel during Robert's visit? (20% and 28%)

Scenario 4, Low

... what do you think is the best single estimate for the likelihood that West Texas State will win the competition? (5% and 25%)

Scenario 4, High

... what do you think is the best single estimate for the likelihood that Harvard will win the competition? (40% and 60%)

Scenario 5, Low

... what do you think is the best single estimate for the likelihood that Austin, Texas, will experience an earthquake strong enough to be life threatening in 1999? (5% and 20%)

Scenario 5, High

... what do you think is the best single estimate for the likelihood that Los Angeles will experience an earthquake strong enough to be life threatening in 1999? (42% and 58%)

Scenario 6, Low

... what do you think would be the best single estimate for the likelihood that it will snow in Leavenworth (Kansas) on the day of the event? (40% and 45%)

Scenario 6, High

... what do you think would be the best single estimate for the likelihood that it will snow at least 1 inch in Moorhead (Minnesota) on the day of the event? (65% and 75%)

Scenario 7, Low

... what do you think would be the best single estimate that a person who was arrested for having 100 or more overdue parking tickets would commit murder some time in their life? (3% and 10%)

Scenario 7, High

... what do you think would be the best single estimate that a drug addict who was arrested for dealing drugs would commit murder within the 3-year period after being released? (35% and 45%)

Scenario 8, Low

... what do you think would be the best single estimate that Peter would be bitten by a snake on his trip? [Peter was described as planning to go to the Swiss Alps for a 2-week tour.] (2% and 8%)

Scenario 8, High

... what do you think would be the best single estimate that Peter would be bitten by a snake on his trip? [Peter was described as planning to go to an Amazon rain forest for a 1-week tour.] (25% and 30%)

Scenario 9, Low

... what do you think would be the best single estimate that more than 10 people will die from heat stroke in Indiana in the summer of 1999? (15% and 30%)

Scenario 9, High

... what do you think would be the best single estimate that more than 20 people will die from heat stroke in Arizona in the summer of 1999? (45% and 60%)

Scenario 10, Low

... what do you think would be the best single estimate that within a 1-month time span—say March of next year—that a homicide will be committed in a national chain bookstore somewhere in the United States? (1% and 9%)

Scenario 10, High

... what do you think would be the best single estimate that within a 1-month time span—say March of next year—that a homicide will be committed in a post office somewhere in the United States? (5% and 20%)

Received May 23, 1997

Revision received May 26, 1999

Accepted June 3, 1999 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.