Why is $p = .90$ better than $p = .70$? Preference for definitive predictions by lay consumers of probability judgments

GIDEON KEREN

Eindhoven University of Technology, Eindhoven, The Netherlands

and

KARL HALVOR TEIGEN

University of Tromsø, Tromsø, Norway

What do people regard as an informative and valuable probability statement? This article reports four experiments that show participants to have a clear preference for more extreme and higher probabilities over less extreme and lower ones. This pattern emerged in Experiment 1, in which no context was provided, and was further explored in Experiment 2 within a positive and a negative context. The findings were further confirmed in Experiment 3, which employed a Bayesian framework with revisions of opinions. Finally, Experiment 4 showed how preference for high probabilities can lead people to prefer an overconfident to a more well-calibrated (accurate) forecaster. The results are interpreted as manifestations of a *search for definitive predictions* principle, which asserts that high probabilities are preferred to medium ones and often favored over the corresponding complementary low probabilities on the basis of their capacity to predict the occurrence of single outcomes.

Uncertainty is an essential and integral facet of all forms of life and is, in principle, unavoidable. The concept of probability is the major tool by which to analyze, express, and communicate uncertainty. In this paper, we take the perspective of the receiver (or consumer) of a probability statement. An author who is told that there is a 70% chance that a paper submitted to a particular journal will be rejected may regard this as "an interesting piece of information." But what if the rejection probability is said to be higher? Or lower? Will all these probabilities be regarded as equally informative, as long as they accurately reflect the rejection rates?

During the last three decades, there have been several attempts to elaborate on psychological aspects associated with probabilistic statements. The research has revolved around three main issues. First, there is the conceptual question of what is probability (which is the only undefined term in the formal probability calculus) and how should it be interpreted. Should it primarily be taken to reflect relative frequencies, or can we also allow for probabilities

of unique events, reflecting external causal tendencies or internal degrees of belief? This key question has created heated debates among probability theorists since the inception of probability theory (Salmon, 1967, is recommended for an illuminating discussion).

More recently, psychologists have been asking the same question about people's conceptions of probability in daily life. For example, Howell and Burnett (1978) proposed a distinction between internal and external uncertainties, depending on whether the outcomes are controllable or not. Kahneman and Tversky (1982) applied the same terms, respectively, to states of uncertainty that are solely determined by internal reflections of the mind and are thus, by definition, subjective and to states of uncertainty that are attributed to the external world and can supposedly be assessed in more objective ways (e.g., by using frequency distributions). They also offered subdivisions of the internal-external distinction, which primarily refer to the type of evidence employed for advocating a particular probability statement. Teigen (1994) adopted a similar approach but proposed a different subdivision of the external-internal distinction. The question of whether people primarily conceive of probabilities when associated with repeated occurrences or with unique events has been raised by several authors (e.g., Hoffrage & Gigerenzer, 1999; Reeves & Lockhart, 1992).

A second, related but different question concerns the extent to which the daily use of probabilities is congruent with the formal theory of probability. In other words,

This research was partly supported by Grant 114193/330 from the Research Council of Norway to K.H.T., and by an internal grant from the Technical University of Eindhoven to G.K. Correspondence concerning this article should be addressed to G. Keren, Faculty of Technology Management, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, or to K. H. Teigen, Department of Psychology, University of Tromsø, 9037 Tromsø, Norway (e-mail: g.b.keren@tm.tue.nl).

is the intuitive probability calculus of lay people compatible with the formal one? The work initiated by Kahneman and Tversky (e.g., Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974), often referred to as the *heuristics and biases* research program, presented a bulk of experimental findings suggesting that, under a wide range of circumstances, human intuitions are incompatible with the formal probability calculus. Empirical evidence in support of this conclusion has been ever increasing (see Baron, 1994, for an updated summary). Notwithstanding, several researchers (e.g., Cohen, 1981; Gigerenzer, 1991; Lopes, 1991) have challenged the validity and interpretations of the experimental results and have expressed fundamental disagreement with the conclusions reached by the heuristics and biases research program.

The third question alludes to whether people are good probability assessors, or how well they can quantify uncertainty. This issue is of both theoretical and practical importance. Given that probability judgments are means of conveying information about uncertainty (Keren, 1991b), the question remains as to how accurate the conveyed information is. Studies addressing the accuracy issue are commonly classified as *calibration* studies (Keren, 1991b; Lichtenstein, Fischhoff, & Phillips, 1982; Yates, 1994).

The above-mentioned three lines of inquiry are directly related to the formal (normative) probability calculus. The first takes the formal theory for granted and explores the possible meanings and interpretations that can be attached to the only primitive that is not defined within the theory—namely, that of probability. The second investigates whether the intuitive calculus of lay persons is congruent with the formal theory. Finally, the third is occupied with determining the accuracy of probability judgments against a normative yardstick.

In the present article, we propose a research program that differs from previous ones in several respects. In particular, rather than adopting the normative theory as a benchmark, we search for a lay theory of probability similar to lay theories of physics (e.g., McCloskey, 1983) or to implicit personality theories (e.g., Chiu, Hong, & Dweck, 1997). The implicit probability theory we envisage cannot be elicited directly, because people would rarely be able to formulate it explicitly. Notwithstanding, we can confirm the existence of such a lay theory, since, as we show empirically, it manifests itself in different ways in people's judgments.

Our empirical approach also differs from previous research in that, rather than asking people to generate probability judgments, we ask them to evaluate probability judgments made by others and test the types of information they unveil from these judgments. Our participants are thus placed in the role of the "consumer" of probability judgments (Yates, Price, Lee, & Ramirez, 1996), and probabilities are used as the independent (rather than the dependent) variable. Notwithstanding, the way people perceive and evaluate probability judgments may

have implications for how these probabilities are generated. For instance, if high probabilities are preferred to low ones, it may invite overconfident estimates.¹

We start our inquiry regarding the implicit information contained in probability judgments by considering an apparently simple question: Which is better, a probability of .7 or a probability of .9? When this question is posed in the abstract, it may appear rather naive and unanswerable. For instance, the reader may ask: Better for what? In what context, and what (if anything) is being predicted? As a first step to making the above question answerable, we should clarify what can be meant by "better." The term implies that probability statements have one or several functions to fill, in which they can be more or less successful. What are these functions?

Probability statements are used in order to communicate some knowledge about the state of the world and, as such, are carriers of information. Forecasting and the assessment of uncertainty are crucial in numerous fields, such as medical diagnosis, personnel selection, meteorology, and economic and political projections, to mention just a few. In daily life, people use probability statements (numerical or verbal) to communicate the uncertainty associated with their knowledge (e.g., I am almost sure that the unfinished symphony was composed by Schubert) or their assessments of events (e.g., Linda will probably be late). Generally, probability statements fulfill two major roles: (1) providing some *information*, even if partial, about the state of the world and (2) supplying some knowledge that may be *valuable* and useful for determining future actions.

When comparing two probabilities, P1 and P2, the question of which is better may be replaced with two slightly more precise (yet not completely unambiguous) questions: Which one is more *informative*? (i.e., linked to the first function mentioned above) and which one is more *valuable*? (i.e., linked to the second function). Several lines of reasoning may determine the answer to each of these questions.

Following one perspective, the one we believe is most compatible with the formal theory of probability, a probability judgment describes a property or a characteristic of an event in question.² It should describe this characteristic as accurately as possible. Consequently, any two probabilities should be judged equally informative and equally valuable, given that they are equally reliable (in the sense of being accurate). Suppose we have two urns with a large number of red and black balls in each, with a different mixture of the two colors. We draw a sample of 100 observations (with replacement) and find that, in the first urn, the ratio between red and blacks is 80:20 and, in the second urn, it is 60:40. If our sole aim is to *describe* (as closely as possible) the characteristics of the two urns—in particular the characteristic regarding the likelihood to *randomly draw a red (black) ball* from each urn—then (given our current knowledge) the probability of a red ball in the first urn is .8 and in the second .6. Neither of these probabilities is more informative or more

valuable, since both describe, with the same level of accuracy, our current knowledge about the state of the two urns.

The above arguments should be qualified by two reservations. First, strictly speaking, the two samples are not equally reliable, since the standard error of the binomial ($\sqrt{p(1-p)/n}$) is not equal for both cases. Whether lay participants are indeed sensitive to this subtle distinction may be questionable. Moreover, the distinction is of relevance only in the context of statistical inference (and sampling distributions). Second, the above reasoning tacitly assumes that there are only two possible states of the world. If more than two possible outcomes exist (e.g., A, B, and C), then $P(A) = .90$ implies that the remainder (.10) is ambiguous, whereas $P(A) = .10$ implies that the remainder .90 is ambiguous. When more than two possible outcome are assumed, one may argue that the former probability is more informative in that it leaves less room for ambiguity.

According to an alternative outlook, probability statements are mainly used for predictions and as guidance for actions, and thus probabilities are informative and valuable to the extent that they can improve decisions (Savage, 1954). How “valuable” or “informative” are probabilities under this view? Consider again the case of the two urns above and suppose that you are offered the following game. You choose one of the two urns and one of the two colors. A ball is randomly drawn from the chosen urn. If the selected ball has the color of your choice, you win \$100. Under these conditions, one may claim that a probability of .8 is more informative, since it indicates more clearly what will be the outcome of the draw. With regard to valuableness, it seems that, under these conditions, the higher the probability of a particular color, the more valuable it is, because it offers more unambiguous advice about which alternative to select, thus having potential practical consequences.

The above considerations suggest that informativeness and valuableness are not necessarily identical under all circumstances and that the interpretation of both may be contingent on context. Specifically, when people are concerned with making concrete predictions of an event that may either occur or not occur (categorical predictions), they will prefer extreme probabilities to medium probabilities, since a probability judgment would be judged more valuable and more informative according to the amount of uncertainty reduction it offers. Experiments 1 and 2 were designed to test this conjecture and to explore how people assess probability statements on dimensions other than likelihood.

EXPERIMENT 1

The purpose of Experiment 1 was to examine whether people will indeed judge higher probability statements to be more valuable and informative than lower probabilities and to explore the reasons they will give to support their choice. We posed this question to groups of students who can be assumed to be generally familiar

with the concept of numerical probabilities, while possessing little knowledge of probability theory and the debate about its conceptual basis. Our prediction was that, for instance, a p of .9 will be more often preferred to a p of .7 than vice versa, for a variety of reasons, but we had no specific predictions about how often they will be regarded as equally good. Also, the situation may be less clear-cut with probabilities below .5. If we compare $p = .2$ with $p = .4$, will people still prefer the higher one? It could be argued that, in this case, the lower probability is actually the one offering most certainty (namely, a probability of .8 that the particular event will *not* occur) and should, on that account, be preferred to the higher one. If the probabilities are symmetric around .5—for instance, in the case of $p = .4$ and $p = .6$ —this line of reasoning should lead to a judgment of *equally informative* *equally valuable*. Still, if probabilities are evaluated from a practical perspective, as a guide to action, some people may prefer the higher probability to the lower also in this comparison.

In this study, we asked participants to assess the relative values of these probabilities “in general,” without implying any particular context. Even if this question would strike some respondents as underspecified and, hence, difficult to answer, it could facilitate speculations of what probabilities mean to them on a more fundamental level. Furthermore, it may help avoid confusions between preferences for particular *probabilities* and preferences for particular *outcomes* of the event being predicted. (In this context, a *preference* for high probabilities means only that people find such statements somehow better than low-probability statements because they are perceived as more useful, definitive, or informative. It does not imply preferences for a particular action or option.)

Method

Participants. The participants were 125 Dutch students from the University of Nijmegen and 108 Norwegian introductory psychology students from the University of Bergen. Both groups participated in a decision-making experiment that consisted of several tasks, including the one described above. The Dutch students were paid approximately \$6 for participating in the experiment, which lasted for about 45 min. The Norwegian students were given no compensation for participation, which formed a part of a lecture.

Design and Procedure. The following written introduction was presented in Dutch or Norwegian to all participants:

As you know, probability statements are used daily in order to express the uncertainty inherent in the world around us. Examples are abundant: weather forecasts, assessing the probability of passing an exam, forecasting different economic indicators, assessing the probability of success of a medical treatment, and so forth. Probabilities are usually expressed on a scale from 0 to 1 (or alternatively, 0%–100%). The higher a probability, the more likely the event under consideration is to occur. The lower the probability, the less likely the event is to occur.

An important aspect of probabilities is that they provide information. In each of the following pages you will be presented with a pair of probabilities. We ask your opinion as to which of the two probabilities is, according to your judgment, more valuable (informative). Circle the option that you believe is more valuable (informative).

The participants were given two out of four pairs of probability assessments (.20 vs. .40, .40 vs. .60, .50 vs. .70, and .70 vs. .90),

presented in terms of percentages (i.e., 60% instead of .60). Half of the participants were asked to judge for each pair which of the two probabilities was more *informative*; the other half had to judge which of the two probabilities was more *valuable*. For each judgment, the participants were asked to circle either one of the two probabilities or a third option—namely, that the two probabilities were equally informative (valuable). Within each pair, the presentation order of the two probabilities was counterbalanced, as was the order of presentation of the two pairs. The participants were also asked to briefly explain their chosen answers.

Results and Discussion

The pattern of results for the Dutch and the Norwegian samples was the same, and thus all the following analyses were performed on the combined sample. The results are portrayed in Table 1, in which the probability pairs are ordered from low (i.e., .20 vs. .40) to high (i.e., .70 vs. .90) and were analyzed by applying logistic regression models. Two types of analysis were performed: one on the entire data and one on a restricted set in which probability pairs that were judged to be equally valuable or equally informative were excluded. Since both the analyses yielded identical results, we report only the analyses on the entire data.

We start with separate analyses of the conditions of valuable and informative. As can be seen, for each and every pair, the higher probability was considered more valuable than the lower one. Similarly, higher probabilities were also judged to be more informative, except for the 20% versus 40% pair. Furthermore, the tendency to prefer the higher probabilities increased with increasing probability level. The higher the values of a probability pair, the more likely was the larger of the two to be judged more valuable ($z = 4.91, p < .001$) or more informative ($z = 3.56, p < .001$). Combining the data for valuable and informative judgments, we obtain a strong tendency to prefer the higher probabilities as the size of probability pairs increases ($z = 6.01, p < .001$) and a significant interaction between this effect and the judgment of valuable-ness or informativeness ($z = 4.45, p < .001$). In other words, as is clearly shown in Figure 1, there was a systematic inclination to prefer the higher probability as the

values of a probability pair increased, and this tendency was significantly stronger in the valuable (compared with the informativeness) condition.

The likelihood that the participants would judge the two probabilities in a given pair to be equal was examined in a separate analysis. The analysis revealed that the smaller a probability pair, the more likely it was to be judged as equal ($z = 2.13, p = .021$) and that the participants in the informative condition, as compared with the participants in the valuable condition, were more likely to judge a pair as equal ($z = 3.65, p < .001$). This tendency was also revealed in within-subjects comparisons, as each participant rated two different probability pairs. For instance, one participant comparing .20 and .40 claimed that both probabilities were equally informative, offering the following reason: "I feel that the probability has little to do with how informative it is. Probability and informativeness are two different things." Yet, when the same subject was asked, on the following page, to compare .70 and .90, he changed his mind and selected .90, commenting apologetically: "Here I have to humble myself; when talking about probabilities as high as 90%, it feels undeniably more informative and trustworthy [than a lower probability]."

High probabilities were preferred to lower for a variety of reasons, the most common being that they offered more certainty and facilitated decisions. The results show that respondents disagreed about probabilities below .50, where some felt that the higher probability, even in this case, offered more certainty, whereas others argued that the lower probability offered more certainty by being farther away from 50/50 (or closer to 0). Lower probabilities were occasionally considered more informative than higher (e.g., .5 was claimed to be more informative than .7), when they appeared easier to understand and interpret.

In a substantial number of cases (about one third of the answers in the valuable condition, and nearly half of all the answers in the informative condition), both probabilities were considered to be equally good. In these cases, most respondents seemed to assert that all probability statements, if correct, are equally informative. This was also offered as a reason for considering them equally valuable. In this condition, the most frequent reason was a different one—namely, that one cannot be preferred to the other as long as we do not know what they refer to. This out-of-context argument was used as a reason for 42 of 79 (53.2%) *equally valuable* judgments, as against 17 of 118 (14.3%) *equally informative* judgments. Some participants went a step further to specify what kind of context they thought would make a difference. Fourteen respondents (17.7%) in the *equally valuable* category said it would matter to know whether the probabilities concerned something positive or negative. This factor was mentioned by only 2 (1.7%) of the respondents giving *equally informative* answers. The effect of positive versus negative contexts is examined in Experiment 2.

In sum, the results show that the simple question of which probability is better is solved by respondents in a variety of ways, reflecting different approaches to the

Table 1
Number of Participants (With Percentages)
Who Chose Each Option in Experiment 1

Response Options	More Valuable		More Informative	
	No.	%	No.	%
20%	17	24.3	18	36.0
40%	26	37.1	10	20.0
Equal	27	38.6	22	44.0
40%	1	2.2	2	2.9
60%	24	52.2	12	17.6
Equal	21	45.7	54	79.4
50%	10	14.9	10	22.2
70%	40	59.7	21	46.7
Equal	17	25.4	14	31.1
70%	0	0.0	2	2.8
90%	32	69.6	42	58.3
Equal	14	30.4	28	38.9

Figure 1. Proportion of participants in the *valuable* and *informative* conditions preferring high over low probability statements, for the four probability pairs employed in Experiment 1.

concept of probability and its use in daily life. We may distinguish at least four lines of thought. (1) One probability may be preferred to another because it simply is higher (closer to 1.0). (2) It may be preferred because it is more extreme (further away from 50/50). The first of these reasons was most frequently offered in the *valuableness* condition, whereas the second was more common in the *informativeness* condition. (3) Alternatively, both probabilities may be considered equal because all probabilities reflect reality (*informativeness* condition), or (4) they may be considered equal because no context is specified (*valuableness* condition).

EXPERIMENT 2

Experiment 1 was aimed to test lay people's evaluation regarding the *valuableness* and *informativeness* of probability statements in the absence of any specified context. Although this abstract form may have induced some vagueness, it was our intention to probe whether there are any fundamental properties of *valuableness* and *informativeness* that cannot be attributed to the thematic content. The orderly pattern of the results suggests that this is indeed the case.

To test the robustness of our results and further explore its implications, a second experiment was conducted. This study was designed as a simplified replication of Experiment 1, asking participants to evaluate two complementary probability pairs, .7 versus .9 and .1 versus .3 (the high- and low-probability pairs in Experiment 1 were not complementary and thus not strictly comparable). As in Experiment 1, separate groups were asked which probability, in each pair, can be considered more *valuable* or more *informative*. Experiment 2 was also designed to test the notion, suggested by some participants in the former study, of a difference between positive and negative contexts. This idea was primarily brought up in the *valu-*

able condition, indicating that a *valuable* probability in a positive context may be different from what is considered *valuable* in a negative context. Because such comments were hardly mentioned in the *informativeness* conditions, we expected that context (positive or negative) would affect judgments of *valuableness*, but not those concerning *informativeness*.

Method

Participants. The participants were 105 students from the University of Utrecht, who participated in a decision-making experiment that consisted of several tasks, including the one described below. The participants were paid approximately \$6 for participating in the experiment, which lasted for about 45 min.

Design and Procedure. As in Experiment 1, the participants were presented with a pair of probabilities and had to judge which of the two was more *informative* (*valuable*) or whether they were equally *informative* (*valuable*). Half of the participants were presented with probabilities associated with positive events. The instructions were identical to those used in Experiment 1, except for the first two sentences, which were replaced by the following:

Probability statements are used daily in order to express the uncertainty inherent in the world around us. Specifically, we are interested here in predictions of positive and desired outcomes. Examples are the probability of having nice weather, assessing the probability of passing an exam, assessing the probability of success of a medical treatment, and so forth.

The other half of the participants were given the same instructions, except that the second and third sentences were replaced by the following:

Specifically, we are interested here in predictions of negative and undesired outcomes. Examples are the probability of having bad weather, assessing the probability of failing an exam, assessing the probability of failure of a medical treatment, and so forth.

Half of the participants in each condition were asked to judge which of the two probability statements was more *valuable*, and the other half were asked which was more *informative*. In each of the four (between-subjects) conditions, the participants were asked to judge the pair 70% versus 90% and the complementary pair 30% versus 10%.

The experiment consisted of several unrelated tasks, including the one described here. The participants were asked to make one of

Table 2
Number of Participants (With Percentages) Who Chose Each Option in Experiment 2,
for Positive and Negative Outcomes, and Comparable Results From a JDM Survey

<i>P</i>	Positive Outcomes				Negative Outcomes				JDM Survey			
	More Valuable		More Informative		More Valuable		More informative		More Useful		More Informative	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
70%	4	17.4	1	3.6	8	33.3	3	10.0	0		1	2.7
90%	13	56.5	10	35.7	7	29.2	16	53.3	7	20.6	10	27.0
Equal	6	26.1	17	60.7	9	37.5	11	36.7	21	79.4	26	70.3
10%	6	26.1	8	28.6	11	45.8	10	33.3	10	34.5	8	27.6
30%	14	60.9	2	7.1	6	25.0	5	16.7	0		0	
Equal	3	13.0	18	64.3	7	29.2	15	50.0	19	65.5	21	73.4

the two judgments at the beginning of the experiment and the other toward the end. Both the order in which the pairs were presented and the order of the presentation of the two probabilities within a given pair were counterbalanced.

Results

The results are shown in Table 2. The data were analyzed in two different ways, employing logistic regression models. As in Experiment 1, all the analyses reported were performed on the entire data, including *same* judgments.

The first analysis was limited to the *positive context* conditions. Given that Experiment 1 did not specify context, we conjectured that the participants assumed a positive context by default, in which case the positive conditions of Experiment 2 can be considered as a replication of Experiment 1. As can be seen from Table 2, higher probabilities were indeed considered more valuable for both high-probability (.90 vs. .70) and low-probability (.30 vs. .10) pairs. There is thus no significant difference between the two pairs in preference for higher probabilities. For judgments of informativeness, however, higher probabilities were preferred for the high pair (.90 vs. .70), but not for the low pair (.30 vs. .10). There is thus a significant difference ($z = 3.37, p < .001$) between the two (high vs. low) pairs in the tendency to prefer the higher probability. This differential pattern for the valuableness and informativeness conditions is exactly the same as that obtained in Experiment 1 (see Table 1, comparing the pairs .90/.70 and .40/.20). Further analysis of the *positive context* data revealed that, as in Experiment 1, the proportion of participants judging a pair to be *equally informative* was significantly larger ($z = 3.24, p < .001$) than the corresponding proportion of participants judging a pair to be *equally valuable*. In sum, as was predicted, the pattern of responses in the positive context was highly similar to that obtained in Experiment 1.

A second analysis (across both the negative and the positive contexts) was performed on valuableness and informativeness judgments separately, in order to test possible context (negative/positive) effects. For judgments of valuableness, higher probabilities were preferred for both the high (.90 vs. .70) and the low (.30 vs. .10) pairs in the positive context, whereas in the negative context, this holds only for the high-probability pair. There was thus a significant interaction ($z = 2.64, p < .01$) between

pair (high or low) and context (positive and negative). No such interaction was found for informativeness judgments. In other words, judgments of informativeness were unaffected by whether the context was positive or negative. Context had, however, an effect on judgments of valuableness—in particular, for the small (probabilities) pair: In the positive context, the majority (14 vs. 6) preferred the higher ($p = .30$) over the more extreme ($p = .10$) probability; in the negative context, this preference was reversed, and the majority (10 vs. 5) of the participants preferred the extreme over the higher probability. We conjecture that the reason for this last finding is that a probability of .10 in the negative context coincides with the desire to avoid the negative outcome and, hence, is perceived as more valuable.

Finally, as in Experiment 1, the proportion of *equal* responses to the informativeness question was significantly higher ($p < .01$), as compared with the valuableness question.

Experiment 2 confirms that there is a difference between the informativeness and the value of probability statements, at least in a positive context. The similarity between Experiment 1 and the positive outcome condition in the present study may indicate that when respondents are asked to evaluate probabilities without context, they provide a predominantly positive context of their own. This is in line with research on the so-called *positivity bias* in judgment and impression formation, which shows that the default option in evaluation is a positive rather than a negative one (Peeters, 1971). The difference between the negative and the positive conditions further indicates that when people are asked which probability is more valuable, they tend to think in terms of which outcome they prefer. For some respondents, *more valuable* may simply be taken to mean *more promising*—that is, they may have interpreted the valuableness question as a question of outcome preference.³

Discussion

The results of the two experiments portray an unequivocal overall disposition for preferring probabilities closer to the extremes, a tendency that is moderated in the lower extreme by a preference for higher probabilities over lower ones. Consequently, the preference for probabilities close to 1.0 is even more pronounced, as compared

with probabilities close to 0. People also seem to differentiate between the informativeness and the valuable-ness of probability judgments. Specifically, the overall proportion (52% and 53% in Experiments 1 and 2, respectively) of participants judging the probabilities of a pair to be *equally informative* was significantly larger than the proportion of participants (34% and 27%, respectively) judging these probabilities to be *equally valuable*.

We have noted already that the question of which probability is more informative or more valuable has no single normative answer and have discussed the different arguments in support of one or the other response. Nonetheless, it is of interest to compare our results obtained from lay people with those from experts. For this purpose, we conducted (through e-mail) a survey among a sample of members from the Judgment and Decision-Making (JDM) society, using exactly the same question format as in Experiment 2, with the exception that the specification of whether the outcome was positive or negative was deleted. Altogether, responses were received from 129 members, the number of respondents in each of the four conditions ranging between 29 and 37.⁴

The results from the JDM survey, shown in Table 2, suggest two major differences between the pattern of evaluations of lay people and experts. First, unlike the lay participants, the preference of JDM members for the extremes is completely symmetrical: In accordance with normative considerations, complementary probabilities (i.e., .90 and .10) are considered to carry the same informational value (and hence are also *equally valuable*). Second, a significantly larger proportion of JDM members did not prefer one probability over the other (70% and 71% for valuable and informative, respectively). Indeed, a central assumption underlying the lay theory of probability is that probabilities are not comprehended in an impartial manner. Rather, lay people tend to interpret probabilities by either attaching or deriving additional information, and this tendency is even stronger when associated with utility (i.e., *valuableness*) than with pure knowledge (i.e., *informativeness*).

Several of the JDM respondents declined to choose an answer, arguing that the question was too abstract and ambiguous. One claim, echoed by several of them, was that it is impossible to judge the informativeness or valuable-ness of a probability judgment without knowing the priors. For example, one respondent wrote, "I like to think of 'informativeness' as depending on the *change* in the probability in the sense that if my (implicit or formal) prior probability for your event had been .89, then learning that $p = .70$ is more 'informative' than learning that $p = .90$." Some respondents linked this idea explicitly to Bayesian principles of probability revision.

Indeed, to be informative, a statement should tell one something new or more than what is already known. This observation forms a basic axiom of Shannon's normative information theory (see Attneave, 1959) and is also reflected in descriptive studies of the perceived information value of verbal statements (Teigen, 1985). From this perspective, the most informative posterior probability

is the one that departs most from the person's prior probability—in other words, the one that requires the largest probability revision.

In Experiment 3, we make the concept of probability revision explicit by stating the prior probability as well as the (posterior) probability that was to be judged for informativeness. In this study, all probability revisions are of the same magnitude. This allows a direct comparison between upward and downward probability revisions of both high (above .5) and low (below .5) probabilities.

EXPERIMENT 3A

Following the Bayesian point of view, probability statements carry information insofar as they allow us to revise our prior opinions about the chances of occurrence of the target event. This idea was brought up by several JDM members in our expert sample but was never mentioned as a reason for informativeness in the lay sample of Experiments 1 and 2. Supposedly, these latter participants did not see prior probabilities as relevant for judging informativeness, or alternatively, they based their judgments on an implicit prior probability of .5. In this case, $p = .9$ may be preferred to $p = .7$ not because it is higher, but because it implies a larger revision of probabilities. To rule out the latter interpretation, we examined whether people would continue to prefer more extreme to less extreme probabilities and high probabilities to low ones in situations in which the prior probabilities are controlled for and the size of probability revision is kept constant.

In Experiment 3A, the participants were asked to compare the informativeness of (posterior) probability statements after the prior probabilities were made known. In all the conditions, there was an absolute difference of .2 between the prior and the posterior probabilities, thus keeping the magnitude of probability revision constant. Our purpose was to test whether the advantage of more extreme versus less extreme probabilities and of high versus low probabilities (comparable with the pattern of preferences found in the previous two experiments) would be maintained.

Method

Participants. The participants were 120 students from the University of Utrecht, who volunteered to participate in this and other decision-making tasks, for which they were paid 11 Dutch guilders (approximately \$6).

Materials. The participants were presented with a scenario of a person named Mr. Peters, who suspects that he has an ulcer and decides to consult a physician (see the example in the Appendix). On the basis of a thorough check-up, the physician makes an initial probability estimate (the *priors*) that Mr. Peters indeed has an ulcer and subsequently sends him for elaborate laboratory tests. Following the laboratory test results, the physician updates his probability estimates (the *posteriors*) upward or downward. A second scenario, presented on the same page, was identical to the previous one, except that it referred to a different person, named Mr. Landman, for whom both the prior and the posterior probabilities were different. After reading the two scenarios, the participants were asked to judge "which of the two patients received more information as a re-

sult of the second estimate (compared with the initial estimate).” Judging that both Mr. Peters and Mr. Landman received an equal amount of information (the correct answer according to Bayesian analysis) was also presented as an option.

Design and Procedure. Four pairs of probabilities (in which the first served as a prior and the second as a posterior) were used: (.20, .40), (.40, .20), (.60, .80), and (.80, .60). The information transmitted by the updated (posterior) probability judgments can be assessed by using Bayes’s theorem, according to which the posterior odds are defined as the product of the prior odds and the likelihood ratios. The likelihood ratio can thus serve as a measure of the additional information transmitted by the posterior probability and, for all the above pairs, is exactly the same.

Four combinations of the probability pairs were inserted in the two scenarios, thus creating four conditions, as shown in the first two columns of Table 3. The participants were randomly assigned to one of the four combinations.

Results

The results are presented in Table 3. As can be seen, 50% of the participants judged the amount of information in the two scenarios to be the same, in accordance with Bayes’s theorem. The remainder of the participants, however, exhibited the same pattern of results as that observed in the previous experiments. In all the four conditions, there was a strong preference for revisions toward the extremes (i.e., 0 or 1.0) or, in other words, for resolution of uncertainty.

Experiment 3A confirms the observations we obtained in the previous experiments—namely, that people have a preference for extreme probabilities—and extends it to a Bayesian framework with revision of opinions. The experiment, however, does not permit testing the asymmetry between low and high probabilities obtained in the previous experiments. The alleged asymmetry was tested in Experiment 3B by comparing the revision of opinion from .40 to .20 with a revision from .60 to .80.

EXPERIMENT 3B

Method

Participants. The participants were 78 students from the Free University of Amsterdam, who were approached on the campus and were willing to contribute a few minutes of their time to complete the task.

Materials and Procedure. The materials and procedure were identical to those used in Experiment 3A, with the exception that only two probability pairs—namely, (.40, .20) and (.60, .80)—were used.

Results

The results are presented in Table 3. Similar to the results of Experiment 3A, more than 50% of the participants judged the two pairs of probability to be equally informative. Of the remaining participants, however, there was a substantial and highly significant ($p < .001$ by a sign test) number who believed that the revision in the upward direction (of $p = 1.0$) was more informative than the comparable revision in the downward direction (of $p = 0$). Apparently, the results obtained in Experiment 3 replicate the results obtained in the first two experiments and extend them to a revision of opinion task.

EXPERIMENT 4

The experiments described above demonstrate the importance people attach to extreme probability judgments—in particular, probabilities close to 1.0. In all the reported experiments, the participants evaluated the relative informativeness or valuableness of two probability statements, which may well have been conceived of as referring to a singular (unique) event. According to the frequentistic view, such probabilities are inadmissible (Salmon, 1967). With repeated (multiple) events, probability statements could refer to the relative frequency of the outcome in question, and it should be easier to agree upon which of the two probability statements is more appropriate. Here, the most informative and most valuable probability statement should be the probability value that most closely mirrors the occurrence frequencies. As was already noted earlier, students of judgment and decision making have stressed the importance of probability accuracy, often termed *calibration*. People who repeatedly say that they are 90% sure, when in fact they are correct only in three out of four cases, are accordingly labeled *overconfident*. It is assumed that confidence statements closer to the actual hit rate (75%) would better reflect reality and thus characterize a *well-calibrated* judge.

Proper calibration may, however, pose a problem for the principle of definite predictions. A well-calibrated forecaster who is correct in 75% of his or her predictions and who, accordingly, claims that each predicted outcome has a $p = .75$ to occur may not be considered sufficiently informative (a single event can either occur or not occur). Many probability consumers (Yates et al., 1996) would prefer to have less ambiguous forecasts—that is, probability statements closer to 1.0. The question posed in the present experiment is whether some people

Table 3
Number of Participants Choosing Each Option
Under the Four Conditions of Experiment 3A and
Under the Single Condition of Experiment 3B

Condition	Mr. Peters	Mr. Landman	
Experiment 3A			
1	$\frac{20 \rightarrow 40}{3}$	$\frac{40 \rightarrow 20}{10}$	$\frac{\text{equally informative}}{17}$
2	$\frac{80 \rightarrow 60}{2}$	$\frac{60 \rightarrow 80}{9}$	$\frac{\text{equally informative}}{19}$
3	$\frac{40 \rightarrow 20}{18}$	$\frac{80 \rightarrow 60}{0}$	$\frac{\text{equally informative}}{12}$
4	$\frac{20 \rightarrow 40}{5}$	$\frac{60 \rightarrow 80}{13}$	$\frac{\text{equally informative}}{12}$
Experiment 3B			
	$\frac{40 \rightarrow 20}{9}$	$\frac{60 \rightarrow 80}{24}$	$\frac{\text{equally informative}}{45}$

would prefer such statements even when the actual hit rate is disclosed. Which consideration is more important: definitiveness or accuracy? How would lay people evaluate forecasts when the criteria of extremeness and accuracy conflict? Experiment 4 was designed to probe these issues.

Method

Participants. The participants were 202 students attending introductory lectures in psychology and philosophy at the Universities of Bergen and Tromsø.

Material and Procedure. The participants were posed the following questions:

Imagine two meteorologists, A and B. Below, you will find a selection of their latest forecasts regarding the probabilities of rainy (sunny) weather. The two weather forecasters participate in a competition about who gives the better forecasts (the winner will receive a prize of NOK 5,000). Please examine carefully their forecasts and tell us who in your opinion is a better forecaster, A or B.

	Forecaster A	Forecaster B
Day 1	90%	70%
Day 2	90%	70%
Day 3	90%	70%
Day 4	90%	70%

In fact, it rained (was sunny) on 3 out of 4 days—in other words, in 75% of the days. Given this information, choose one of the following options:

- A should win the prize
- B should win the prize

Give a brief explanation for your choice.

About half of the participants responded to forecasts of rain, whereas the other half responded to forecasts of a sunny day. In both cases, half of the questionnaires did not specify on which particular days the weather was actually nice or rainy (as in the above example), whereas the other half specified the “actual outcomes” day by day, with the deviant outcome listed on Day 3. There were no differences between these two versions of the questionnaires, so the results were pooled.

The participants in the first condition ($n = 82$) received scenarios in which the forecasters predicted sun or rain with probabilities of 90% and 70%, respectively, as is shown in the above example. The participants in the second condition ($n = 80$) received identical scenarios, except that the complementary probabilities (i.e., 10% and 30%) were used. Here, the actual outcomes were specified as sun (rain) on 1 out of 4 days (i.e., in 25% of the cases). In an additional condition, a separate group ($n = 40$) was given the same problem as in the first condition, except that B's probabilities were changed from 70% to 75%.

Following normative considerations, the forecasts of both meteorologists portray the same resolution (namely, 0), but B's forecasts are conclusively better calibrated. In the third condition, he is actually perfectly calibrated, reporting probabilities of 75% when he is, in fact, correct on 75% of the days. If the two forecasters are judged exclusively according to criteria of accuracy, forecaster B's performance is clearly superior. However, A's forecasts are more extreme and more decisive, and he may be preferred on the ground that he was “closer” to the actual outcome on 3 out of the 4 days.

Results and Discussion

As can be seen from Table 4, over all conditions, the majority of the participants (close to 58%) exhibited preference for extreme probabilities over well-calibrated ones. Apparently, and consistent with the results of the first three experiments, extremeness of probabilities looms over other considerations when assessing the quality of probability

Table 4
Number of Participants (With Percentages) in Experiment 4 Preferring Forecasters With Extreme Versus Moderate Probabilities (for Sunny or Rainy Days) Under Conditions With High Probabilities ($p = .70$ vs. $p = .90$), With Complementary Low Probabilities ($p = .30$ vs. $p = .10$), and With a Comparison of a Perfectly Calibrated Forecaster ($p = .75$) With an Over-Forecaster ($p = .90$)

Condition	Sunny	Rainy	Total	
			No.	%
High probabilities				
Forecaster A: $p = .90$	19	24	43	52.4
Forecaster B: $p = .70$	23	16	39	47.6
Complementary low probabilities				
Forecaster A: $p = .10$	21	28	49	61.0
Forecaster B: $p = .30$	17	14	31	39.0
Comparison with perfectly calibrated forecaster				
Forecaster A: $p = .90$	11	14	25	62.5
Forecaster B: $p = .75$	8	7	15	37.5

judgments. This was even the case in the condition in which we attempted to stress the accuracy dimension by making forecaster B perfectly calibrated. In addition, the instructions specifically mentioned that it rained “on 75% of the days,” highlighting the actual rate of occurrence. Our results are congruent with recent studies reported by Yaniv and Foster (1995), who concluded that “people might accept errors in the interest of securing more informative judgments” (p. 424).

Asked for their reasons, many participants who preferred forecaster A noted that forecaster A made more conclusive predictions, and “came closer” on 3 out of the 4 days. Thus, this forecaster appeared to be “more often more right.” In contrast, they claimed, none of the predictions made by forecaster B was sufficiently decisive. Implicit in the participants' reasoning is the idea that a probability for rain (or sun) of .9 is so close to 1.0 that the decision to take an umbrella (or leave it at home) is unequivocal. In retrospect, that decision was indeed justified in 3 out of the 4 cases, and the single error is thus underweighted. In contrast, a probability of .75 is not sufficiently high for a definitive decision and thus, in each case, creates ambivalence. The “average” accuracy of the forecasts (i.e., calibration) seems to be of little importance. Even among participants who opted for forecaster B, fewer than half referred to average accuracies or closeness to actual hit rates. These participants favored forecaster B mainly because he made the smaller error on the day that both forecasters were wrong or because they appreciated his generally cautious attitude.

The preference for definitive predictions is particularly applicable when the number of predictions is small (or singular). Naturally, under multiple repeated trials (i.e., say we had 100 forecasts made by each of the two forecasters), it would be impossible for participants to examine each forecast separately. Under such circumstances, they may tend to assess the “average” (expected) performance and most likely prefer the more calibrated forecaster. With few forecasts (as in our experiment), however, what matters is the extent to which a forecast provides a definite and unequivocal prediction.

Indeed, the desire for certainty and the overweighing of probabilities close to 1.0 have been long recognized by students of decision making as being reflected in the so-called *certainty effect* (Kahneman & Tversky, 1979). As in the present research, Keren and Wagenaar (1987; see also Keren, 1991a) demonstrated that, with multiple repeated trials, other considerations may loom that may weaken the certainty effect or even make it vanish altogether.

GENERAL DISCUSSION

In the present article, we have shown that the manner in which lay people interpret and understand probabilistic statements goes beyond what is intended by the formal theory.

The fact that participants use different criteria was demonstrated in the first two experiments, in which people employed two different types of responses. Some participants were inclined to think of probabilities as merely describing states of the world and hence judged different probabilities to be equally informative (or equally valuable), whereas others seemed to adopt a more instrumental attitude, placing emphasis on the predictive potential of probability statements. According to this latter view, probabilities close to the extreme are preferred to medium probabilities, and high probabilities are preferred to low ones. This is in line with what we have termed the *principle of definitive predictions*: Only extreme probabilities provide a distinct prediction of what will happen next. (When predicting outcomes of repeated trials, a case can be made for saying that, for example, $p = .7$ gives an equally definite prediction of what will happen on the average, or in the long run, as does a probability of .9.) Participants judging which probability was more valuable (rather than more informative) showed an even stronger preference for extreme probabilities, especially probabilities close to 1.0. Our interpretation is that the participants in these conditions were even more encouraged to view probabilities as a tool for making exact predictions and hence favored those probability statements that more strongly indicated what to expect.

One strength—and at the same time, a weakness—of the first two experiments was the lack of any specific context description. This could have led participants to provide their own context. They may have also preferred extreme probabilities simply because these probabilities indicated a greater revision of an assumed (default) prior probability of .50 and hence would be more informative than probabilities that required a smaller revision. This hypothesis was tested in Experiment 3, in which an explicit context was provided and the magnitude of probability revision was kept constant.

Experiment 4 offered further evidence for the preference of extreme probabilities. For a substantial number of participants, this preference overrode considerations of long-run accuracy. The important point to notice here is that the pattern of results, as well as the participants'

comments, suggests that the participants evaluated each of the four predictions separately, rather than being concerned about the expected accuracy of the multiple events. Indeed, in daily life, this is often the case. For instance, casual observations suggest that individual patients expect their physician to provide them with an unqualified diagnosis and unambiguous course of action. The physician's "average" performance would be of less interest.

These results carry important implications for the interpretation of calibration studies. First, they question the importance that many researchers in this field have attributed to the accuracy of probability judgments (Keren, 1997). Evidently, and in agreement with Yates et al. (1996), people may (at least under some circumstances) be more concerned about the resolution of probabilities, thus overweighing events with high probabilities that actually materialized. Second (and related), as is suggested by the results of Experiment 4, when evaluating probabilities, people attend not just to the average prediction accuracy, but even more so to single outcomes. What people value most is the ability to discriminate between occurrence ($p = 1.0$) and nonoccurrence ($p = 0$), an ability that, in the calibration research, is termed *resolution*. Finally, an important implication of the experiments is concerned with the pervasive finding that people tend to be overconfident in their probability judgments (see, e.g., Keren, 1991b), a finding that has recently become controversial. Notwithstanding that part of the observed overconfidence in many experiments may be due to regression effects (Erev, Wallsten, & Budescu, 1994), the present experiments strongly suggest that overconfidence is not merely an artifact but could reflect people's preference for high probabilities.

Common to all the four experiments is the underlying desire for certainty that is embedded in the principle of the search for definitive predictions. This principle can alternatively be viewed as uncertainty avoidance. There are several possible accounts for why people dislike uncertainty and thus prefer probabilities in the extreme, to an extent that sometimes appears to dominate other considerations. First, uncertainty (in particular, when associated with possible negative outcomes) often entails an unpleasant and undesirable affection that people want to alleviate. A convincing example is offered by Sieff, Dawes, and Loewenstein (1997), who conducted a study on the stress associated with being diagnosed as HIV positive. Their results show that the actual experienced distress reported by participants who were diagnosed as being HIV positive was apparently lower than the distress these same participants anticipated before being tested. Supposedly, the relief obtained from resolving the uncertainty somewhat lowered the stress for many of these people, despite their being identified as HIV positive. The point to be emphasized is that uncertainty, in and by itself, often entails an undesired emotional burden that one wants to get rid of.

Besides the emotional facet associated with uncertainty, there is also a cognitive component related to cu-

riosity. Recent research interprets curiosity as a form of cognitively induced deprivation (Loewenstein, 1996). The need for uncertainty resolution may thus reflect the desire to satisfy this deprivation.

Finally, uncertainty may be avoided because it is seen as an obstacle for efficient decision making. As was already noted in the discussion of Experiment 4, extreme probabilities may be desired because they facilitate decisions regarding actions and minimize potential internal conflicts. The contrast between the concepts of *uncertainty* and *decision* was demonstrated in a semantic differential study in which the connotations of these two concepts were explored (Teigen, 1996). Whereas *decision* received particularly high ratings as strong, active, certain, positive, stable, and wise, *uncertainty* was rated as weak, passive, negative, unstable, and stupid. Uncertainty is accordingly viewed as dysfunctional and counterproductive in a context in which action and decisiveness are required.

The present results should not lead us to believe that high probabilities will universally be preferred to lower or intermediate probabilities, even if the outcome in question should actually occur. A new set of studies (Keren & Teigen, in press) shows that high probabilities are also believed to manifest themselves as high-amplitude or strong outcomes. Thus, a soccer team that has a high probability of winning is also believed to score earlier and win with a larger margin than does a team that is assigned a lower probability of winning. An implication of this is that a probability of .7 may occasionally be seen as more "correct" than a probability of .9, even for a winning team, when the team in question wins with a small margin. Thus, the principle of definitive predictions, where large probabilities are seen as more appropriate than lower probabilities, applies primarily to outcomes that can be described categorically as present or absent (occurrences vs. nonoccurrences). When outcomes are described in quantitative terms, such as a strong (5 on the Richter scale) or a weak (1.5 on the Richter scale) earthquake, or a large (6:1) or a small (1:0) win in a soccer game, people expect these differences to be reflected in probability judgments. In a recent set of experiments, we have shown that people prefer low probabilities to reflect "weak" or "small" outcomes and high probabilities to reflect "strong" or "large" outcomes and have labeled this tendency as the probability-outcome correspondence principle. Under some circumstances, this principle is incompatible with the principle of definitive predictions explored in the present paper.

REFERENCES

- ATTNEAVE, F. (1959). *Applications of information theory to psychology*. New York: Holt, Rinehart & Winston.
- BARON, J. (1994). *Thinking and deciding*. Cambridge: Cambridge University Press.
- CHIU, C., HONG, Y., & DWECK, C. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality & Social Psychology*, **73**, 19-30.
- COHEN, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral & Brain Sciences*, **4**, 317-331.
- EREV, I., WALLSTEN, T. S., & BUDESCU, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, **101**, 519-527.
- GIGERENZER, G. (1991). How to make cognitive illusions disappear: "Beyond heuristics and biases." In W. Stroebe & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 2, pp. 83-115). Chichester, U.K.: Wiley.
- HOFFRAGE, U., & GIGERENZER, G. (1999). The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, **12**, 275-305.
- HOWELL, W. C., & BURNETT, S. A. (1978). Uncertainty measurement: A cognitive taxonomy. *Organizational Behavior & Human Performance*, **22**, 45-68.
- KAHNEMAN, D., SLOVIC, P., & TVERSKY, A. (EDS.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- KAHNEMAN, D., & TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263-291.
- KAHNEMAN, D., & TVERSKY, A. (1982). Variants of uncertainty. *Cognition*, **11**, 143-157.
- KEREN, G. (1991a). Additional tests of utility theory under unique and repeated conditions. *Journal of Behavioral Decision Making*, **4**, 297-304.
- KEREN, G. (1991b). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica*, **77**, 217-273.
- KEREN, G. (1997). On the calibration of probability judgments: Some critical comments and alternative perspectives. *Journal of Behavioral Decision Making*, **10**, 269-278.
- KEREN, G., & TEIGEN, K. H. (in press). Probability-outcome correspondence in lay interpretations of probability statements. *Memory & Cognition*.
- KEREN, G., & WAGENAAR, W. A. (1987). Violations of utility theory in unique and repeated gambles. *Experimental Psychology: Learning, Memory, & Cognition*, **13**, 387-391.
- LICHTENSTEIN, S., FISCHHOFF, B., & PHILLIPS, L. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). Cambridge: Cambridge University Press.
- LOEWENSTEIN, G. (1996). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, **116**, 75-98.
- LOPES, L. (1991). The rhetoric of irrationality. *Theory & Psychology*, **1**, 65-82.
- MCCLOSKEY, M. (1983, February). Intuitive physics. *Scientific American*, pp. 122-130.
- PEETERS, G. (1971). The positive-negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, **22**, 571-591.
- REEVES, T., & LOCKHART, R. S. (1992). Distributional versus singular approaches to probability and errors in probabilistic reasoning. *Journal of Experimental Psychology: General*, **122**, 207-226.
- SALMON, W. C. (1967). *The foundation of scientific inference*. Pittsburgh: University of Pittsburgh Press.
- SAVAGE, L. J. (1954). *The foundations of statistics*. New York: Wiley.
- SIEFF, E. M., DAWES, R. M., & LOEWENSTEIN, G. (1997). *Anticipated versus actual reaction to HIV test results* [Working paper]. Pittsburgh: Carnegie Mellon University.
- TEIGEN, K. H. (1985). Perceived informativeness of verbal information. *Current Psychological Research & Reviews*, **4**, 3-16.
- TEIGEN, K. H. (1994). Variants of subjective probabilities: Concepts, norms, and biases. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 211-238). Chichester, U.K.: Wiley.
- TEIGEN, K. H. (1996). Decision making in two worlds. *Organizational Behavior & Human Decision Processes*, **65**, 249-251.
- TVERSKY, A., & KAHNEMAN, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, **185**, 1124-1131.
- YANIV, I., & FOSTER, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General*, **124**, 424-432.
- YATES, J. F. (1994). Subjective probability accuracy analysis. In G. Wright &

- P. Ayton (Eds.), *Subjective probability* (pp. 381-410). Chichester, U.K.: Wiley.
- YATES, J. F., PRICE, P. C., LEE, J. W., & RAMIREZ, J. (1996). Good probabilistic forecasters: The "consumer's" perspective. *International Journal of Forecasting*, 12, 41-56.

NOTES

1. Being a rather unexplored topic, the question of probability preferences does not allow us to run experiments designed to test competing theories. Instead, we are using an inductive approach, asking participants which probabilities they prefer, which leads us to the theoretically important view that people first and foremost appreciate probabilities that allow them to anticipate concrete (singular) outcomes and will, accordingly, have a preference for more extreme and higher probabilities over less extreme and lower ones.

2. This perspective is based on attributing uncertainty to the external world, in which a system is conceived as containing different dispositions to produce different events (e.g., Kahneman & Tversky, 1982). Notice

that this view is compatible with both the frequentistic and the subjective interpretations of probability.

3. One may think that almost everyone would consider higher probabilities of positive outcomes and lower probabilities of negative outcomes to be more valuable. Evidently, this is not entirely true. For instance, for positive outcomes, about 26% believe a probability of .10 to be more valuable than a probability of .30. Supposedly, because .10 is more definitive and both .10 and .30 are low probabilities, the former has the advantage of not building up unwarranted expectations, thus averting unnecessary disappointment.

4. Half of the participants in the JDM sample were asked which probability statement is "more informative," whereas the other half were asked which statement they would consider "more useful." The reason for using the latter term (rather than "more valuable" as in Experiments 1 and 2) was to emphasize even more strongly the practical implications of probability statements. Although several respondents (in both groups) commented that extreme probabilities would be more important in a practical than in a purely theoretical context, there was no difference in the preference pattern of the two groups, as is shown in Table 2.

APPENDIX

Instructions to Participants in Experiment 3

Consider the following two events.

Mr. Peters has suffered from stomach ache for the last few days. He suspects that he may have ulcer and decides to visit the physician. Following a comprehensive check-up, the physician estimates the chance that Mr. Peters has an ulcer at 20%. He further sends Mr. Peters for elaborate laboratory tests. Two days later, Mr. Peters visits the physician again. The physician has meanwhile received the laboratory results. On the basis of these results, he now estimates the probability that Mr. Peters has an ulcer to be 40%. In other words, the physician has changed his estimate that Mr. Peters has an ulcer from 20% to 40%.

Mr. Landman has suffered from stomach ache for the last few days. He is afraid that he has an ulcer and decides to visit the doctor. Following a comprehensive check-up, the doctor estimates the chance that Mr. Landman has an ulcer at 40%. He further sends Mr. Landman for elaborate laboratory tests. Two days later, Mr. Landman visits the doctor again. The doctor has meanwhile received the laboratory results. On the basis of these results, he now estimates the probability that Mr. Landman has

an ulcer to be 20%. In other words, the physician has changed his estimate that Mr. Landman has an ulcer from 40% to 20%.

In both cases, the physician has changed his probability estimate. Moreover, in both cases, the second estimate provides new information. The question we would like you to answer is as follows: Which of the two patients received more information as a result of the second estimate (as compared with the initial estimate). Take a few moments for thought and circle the option that best represents your opinion. Note: There is no "right" or "wrong" answer. We are simply interested in your opinion.

Mr. Peters received more information, because a modification from 20% to 40% offers more information than a modification from 40% to 20%.

Mr. Landman received more information, because a modification from 40% to 20% offers more information than a modification from 20% to 40%.

Both Mr. Peters and Mr. Landman received the same amount of information.