Journal of Experimental Psychology: General
2006, Vol. 135, No. 3, 429–447

0096-3445/06/$12.00 DOI: 10.1037/0096-3445.135.3.429

Inflation of Conditional Predictions

Asher Koriat
University of Haifa

Klaus Fiedler
University of Heidelberg

Robert A. Bjork
University of California, Los Angeles

The authors report 7 experiments indicating that conditional predictions—the assessed probability that a certain outcome will occur given a certain condition—tend to be markedly inflated. The results suggest that this inflation derives in part from backward activation in which the target outcome highlights aspects of the condition that are consistent with that outcome, thus supporting the plausibility of that outcome. One consequence of this process is that alternative outcomes are not conceived to compete as fully as they should. Another consequence is that prediction inflation is resistant to manipulations that induce participants to consider alternative outcomes to the target outcome.

Keywords: conditional prediction, probability judgment, confirmation bias, metacognition, overprediction

> A Jewish joke goes like this: Two litigants came before the rabbi. After hearing the first testimony, the rabbi says, "It seems that you are right."
>
> But after the second man speaks, the rabbi says, "It seems that you are right, too." "How can this be?" says the rabbi's wife, who has been listening to the arguments. "How can both of these men be right?" "Hm," says the rabbi. "You know what? You're also right."
>
> —*The Big Book of Jewish Humor*

In many real-life situations people make predictions about the likelihood of a particular outcome given a certain present or future condition. Doctors sometimes need to assess whether the administration of a new treatment to a patient will be beneficial. In deciding whether to release a convict on bail, judges assess the chances that the person in question will stand for trial later. Investors make predictions about the possible consequences of future political or economic developments on the value of particular shares. And, of course, in daily life people routinely make predictions about the possible consequences of various actions that they may take. Such predictions often have important implications for their decisions and behavior (Tversky & Koehler, 1994).

The studies reported in this article focus on conditional predictions, that is, on the perceived probability that a certain outcome will occur given a certain condition. We propose that conditional predictions tend to be markedly inflated. Their inflation derives in part from a backward activation process in which the stated outcome, whose likelihood is to be assessed, brings to the fore aspects of the condition that support the occurrence of that outcome.

In what follows we first review evidence for the general bias toward overestimation of the probability of events under different conditions. Focusing then on conditional predictions, we detail our conceptual framework and outline our predictions.

The Overestimation of the Probability of Future Events

One of the pervasive biases in probability judgments is the tendency to overestimate the likelihood of occurrence of hypothetical target outcomes. Merely specifying a particular future event or outcome leads people to perceive that event or outcome as more likely (Fiedler, 2000; Fiedler, Armbruster, Nickel, Walther, & Asbeck, 1996; Koehler, 1991; Tversky & Koehler, 1994). Such overestimation bias is particularly strong when participants are instructed to imagine or to explain the outcome before judging its likelihood (e.g., Hirt & Markman, 1995). For example, participants who were asked to imagine that Gerald Ford would win the 1976 presidential election judged a Ford victory as more likely than those who imagined a Carter victory (Carroll, 1978). Imagination has been found not only to inflate the probability of future events but also to produce memories for events that had not happened (Garry & Polaschek, 2000).

Explaining the reasons for a possible future event has also been shown to increase the perceived likelihood of that event (Ander-

Asher Koriat, Department of Psychology, University of Haifa, Haifa, Israel; Klaus Fiedler, Institute of Psychology, University of Heidelberg, Heidelberg, Germany; Robert A. Bjork, Department of Psychology, University of California, Los Angeles.

Portions of this work were presented at the meeting of the Psychonomic Society, Toronto, Ontario, Canada, November 12, 2005. By coincidence, William Maki presented at the same session empirical work that is very related, although it was cast within a different conceptual framework that emphasizes memory processes (Maki, 2005b).

The research reported was conducted in the Institute of Information Processing and Decision Making, University of Haifa. We gratefully acknowledge support for this research by the German Federal Ministry of Education and Research (BMBF) within the framework of German-Israeli Project Cooperation (DIP). We thank Rinat Gil for conducting the experiments and Limor Sheffer for her help in the analyses of the data. We are indebted to Ido Erev, Raymond Nickerson, Jim Sherman, and Karl Halvor Teigen for their comments on a draft of this article.

Correspondence concerning this article should be addressed to Asher Koriat, Department of Psychology, University of Haifa, Haifa 31905, Israel. E-mail: akoriat@research.haifa.ac.il

son, Lepper, & Ross, 1980; Hirt & Sherman, 1985; Koehler, 1991). For example, participants who were asked to explain a particular hypothetical outcome of a football game judged that outcome to be more likely than participants who explained the opposite outcome (Sherman, Zehner, Johnson, & Hirt, 1983). In Hoch's (1984) study, participants generated reasons why a future event might and might not occur. A primacy effect was observed such that judgments were influenced more by whatever side of the issue participants thought about first. The solicitation of an explanation has been also found to lead to the perseverance of people's initial theories (Anderson et al., 1980) but soliciting explanations for alternative hypothetical outcomes to the target outcome was found to help debias likelihood judgments (Hirt & Markman, 1995).

Conditional Predictions

In this study we focus on conditional predictions. Conditional predictions are prompted by questions that conform to the following format: "What is the probability that event *y* will occur under condition *x*"? The two terms in this question represent a target outcome (*y*) and an envisioned condition (*x*), respectively. Thus, in soliciting a conditional prediction, a specific condition is mentioned and a specific possible outcome is explicitly stated. The condition can sometimes be influenced by judges themselves (e.g., "how likely is my wife to be angry if I tell her . . .?"), but in other cases it may be externally controlled ("what is the likelihood that if Iran stops supplying crude oil, the price of a barrel will exceed $85?").

We will evaluate the hypothesis that in making conditional predictions, people are subject to a prediction inflation bias, overestimating the likelihood of occurrence of the stated outcome given the stated envisioned condition. As just noted, participants tend to overestimate the likelihood of target events when they are first asked to imagine or explain these events (Sanbonmatsu, Posavac, & Stasney, 1997). In this study, however, we focus on predictions made in situations in which neither imagination nor explanation of the target outcome are explicitly solicited. In such situations an important mechanism that may foster inflated predictions is backward activation: Merely considering the outcome in conjunction with the stated condition highlights aspects of the condition that might not have been transparent had that outcome not been mentioned. This occurs when the condition is rich enough so that it can submit itself to different constructions. The activated features of the condition tend to be those that support the target outcome, and hence make it seem more plausible or likely than it actually is. Thus, we expect the magnitude of prediction inflation to increase as a function of the number of features of the condition that are activated by a consideration of the potential outcome—features that would not have been considered otherwise.

Because alternative outcomes may bring to the fore different aspects of the condition, each such outcome may appear plausible or likely, with relatively little competition between them. As a consequence, even though one possible outcome may be perceived as very likely given the stated condition, when a second, alternative outcome is presented, it may modify the construal of the condition so as to make that outcome also seem quite likely. Hence the sum total of the perceived likelihood of a set of mutually exclusive outcomes may sometimes exceed one (see Fiedler & Armbruster, 1994; Sanbonmatsu et al., 1997; Teigen, 1983; Tversky & Koehler, 1994).

Confirmation Bias, Hindsight Bias, and Conditional Predictions

The inflation of conditional predictions may have much in common with other phenomena discussed in the literature (see Fiedler et al., 1996), particularly confirmation bias and hindsight bias. Confirmation bias is perhaps the best studied bias in decision-making research (see Nickerson, 1998; Trope & Liberman, 1996). It refers to the tendency to justify a conclusion that has already been reached by selectively utilizing supportive evidence. Confirmation bias has been claimed to underlie the overconfidence that is typically observed when participants are required to indicate their confidence in the correctness of an answer that they have chosen or provided (see Fischhoff, Slovic, & Lichtenstein, 1977; for reviews, see Ayton & McClelland, 1997; Erev, Wallsten, & Budescu, 1994; Nickerson, 1998; but see Winman & Juslin, 2006). The prediction inflation is analogous to a confirmation bias except that it is assumed to operate in making predictions about the future. Our view is similar to that advanced by Koehler (1991) and Fiedler et al. (1996). In discussing the effects of explanation on probability judgments, Koehler proposed that the explanation task draws attention to the specified, focal hypothesis. Once a focal hypothesis is established, the person may be induced to adopt a conditional reference frame in which the focal hypothesis is temporarily assumed to be true, and then assess how that hypothesis can plausibly account for the relevant evidence. Fiedler et al. (see also Fiedler, 2000) argued that the mere considering of a proposition tends to induce the belief that the proposition is true, independent of any confirming evidence.

In a similar manner, we propose that conditional predictions are performed by building a scenario that leads from the condition to the outcome. In doing so, the construal of the stimulus condition is "colored" by the stated outcome in the direction of a greater emphasis on aspects of the condition that are consistent with the specified outcome. Thus, the stated outcome helps to make accessible aspects of the condition that may not come forward under a forecasting mode, in which participants are called upon simply to generate possible outcomes given the specified condition. Conditional predictions, then, are similar to postdictions in which participants consider the features that the condition and outcome have in common.

In this respect, conditional predictions are also similar to the hindsight bias, first demonstrated by Fischhoff (1975; for reviews, see Christensen-Szalanski & Willham, 1991; Hawkins & Hastie, 1990; Hoffrage, Hertwig, & Gigerenzer, 2000; see also Hoffrage & Pohl, 2003, for a journal issue dedicated to the topic). When people are asked to recall their earlier answer to a question that calls for a quantity estimate and later receive the correct answer, their recollection of their original estimate tends to shift toward the correct answer. Similarly, when people predict the outcome of a historical event and are then asked to recollect their prediction after the outcome is known, they tend to increase the postdicted likelihood of that outcome.

The prediction inflation investigated in this study is, in a sense, a mirror image of the hindsight bias. Whereas in the hindsight bias the participant's past predictions are distorted in retrospect once the actual outcome is revealed, in prediction inflation the participant's predictions are distorted in the direction of the stated outcome whose future likelihood must be assessed. One mechanism that has been proposed to underlie the hindsight bias is that when people attempt to reconstruct the cognitive processes that may have led them to arrive at an answer, their reconstruction is biased by the knowledge of the correct answer. In a like manner, in making conditional predictions people may find it difficult to escape the influence of the stated outcome, building a scenario that is biased toward that outcome.

Finally, the process postulated to engender prediction inflation is most closely related to that underlying the foresight bias that has been reported recently by Koriat and Bjork (2005, in press). The foresight bias refers to the tendency of learners to experience an illusion of competence during learning, making inflated predictions about the likelihood of recalling the studied materials at test. According to Koriat and Bjork, learners' predictions are prone to overconfidence because judgments of learning are made in the presence of information that is absent, but solicited, on a subsequent test—such as the targets in cue-target paired associates. Such information may bring to the fore aspects of a cue that will be less apparent when the cue is later presented alone. The experiments we report have much in common conceptually and methodologically with the experiments used to support the foresight bias that occurs during learning.

The Present Study

The task we used in the following experiments (see also Garskof & Forrester, 1966; Maki, 2005a, 2005b) is one for which normative data are available. These data consist of word association norms that list the percentage of people who give a certain response word (outcome) as the first word that comes to mind in response to a particular stimulus word (condition). Participants in our experiments were essentially required to make a conditional prediction of the form "suppose that a person is presented with word *x* (condition) and asked to respond with the first word that comes to mind, what is the likelihood that he/she will respond with the word *y* (outcome)?" Thus, both the condition and the outcome are specified.

Whereas Experiment 1 sought to establish the inflation of conditional predictions, subsequent experiments had several additional aims. The first aim was to examine the degree of perceived competition between alternative responses to the same cue word. We expected an overprediction effect, whereby the estimated percentages of occurrence of several alternative responses (even only two such responses) to the same cue word may sum up to more than 100%. The second aim was to try to relate the prediction inflation effect to the backward associations from the target word to the cue word. Toward that aim, word pairs with asymmetric associations were used such that the primary association was either from the cue word to the target word (forward) or from the target to the cue (backward). If prediction inflation derives from a backward activation in which the outcome brings to the fore aspects of the condition that are consistent with it, then it should be particularly strong for backward-associated pairs in which the association is from the target word (outcome) to the cue word (condition). The third aim, finally, was to examine the extent to which prediction inflation could be remedied by manipulations that induce participants to consider alternative outcomes other than the one specified. These manipulations were also intended to provide some insight regarding the mechanism underlying prediction inflation.

The word association task used in this study has several features that make it convenient for investigating these issues. The first, just mentioned, is the availability of normative data with which the participants' responses can be compared. The second, is that the semantic-associative properties of the word pairs allowed us to trace prediction inflation to the tendency of the response word to emphasize associative and semantic features of the stimulus word that are less likely to come forward when the response word is not stated (see also Koriat, 1981; Koriat & Bjork, 2005). As noted earlier, we did so by using asymmetrically associated pairs. The third feature, finally, is that participants can practice the task of producing an association to the cue word and thus can have firsthand experience with the behavior that they are trying to predict. In fact, it has been proposed that people generally rely on their own subjective experience in making predictions for others. For example, Kelley and Jacoby's (1996) results suggest that participants judge the difficulty of different anagrams for others by observing their own experience in attempting to solve these anagrams themselves (see also Kelley, 1999). Nickerson (1999), reviewing studies dealing with the question of how one knows what others know, concluded that people begin by imputing their knowledge to others.

If participants' predictions are affected by their own experience producing associations, then asking participants to generate overtly their own responses prior to making predictions about others should reduce prediction inflation. In fact, when participants produce a different response from the target response that they are subsequently asked to judge, we might even expect an underestimation of the occurrence of the target response. However, if the presentation of a stated outcome activates aspects of the condition that are consistent with it, as we have proposed, then generating one's own response may not be effective in eliminating or reducing prediction inflation even when the participant's generated response differs from the stated target response. Thus, it might be the case that after producing one's own response, the presentation of a different target word (outcome) along with the cue word (condition) may modify the representation of the condition to the extent of preempting the experience gained from generating one's own response to the same nominal stimulus. In general, the postulated mutability of the representation of the condition in response to different considered outcomes should make prediction inflation quite resistant to change.

Experiment 1: Prediction Inflation

Experiment 1 evaluated the existence and magnitude of prediction inflation. Participants were given a description of the word association task in which people are presented with a stimulus word (cue) and asked to say the first word (target) that comes to mind. They were then presented with a list of cue−target pairs and asked to estimate for each pair the percentage of participants who would respond with the target word when presented with the cue word.

Method

Participants. Thirty Hebrew-speaking undergraduates participated in the experiment during a class meeting.

Materials. A list of 45 Hebrew word pairs was compiled, including 15 pairs at each of three levels of associative percentage: zero, low, and high. Associative percentage was defined as the percentage of people who gave the second word of the pair (target) as the first response to the first word (cue) in a word association task, based on Hebrew word-association norms (Rubinsten, Anaki, Henik, Drori, & Faran, 2005). Mean percentage of responding for the zero, low, and high word pairs was 0, 4.4 (range = 3.9 to 4.9) and 23.9 (range = 21.6 to 27.5), respectively.

Procedure. The instructions followed closely those used by Rubinsten et al. (2005) in collecting their word association norms. They were as follows (paraphrased from Hebrew):

> In a word association task, a person is presented with a word and is asked to say as fast as he can the first word that comes to mind in response to the presented word.
>
> You will be presented with pairs of words. For each pair we would like you to estimate the percentage of people who would say the second word (on the right) as the first response to the stimulus word (on the left).[1]
>
> You will find a blank line next to each word pair. Write on that line a number between 0 and 100 that reflects the percentage of people who, according to your opinion, would say the second word (the left one) in response to the first word (the right one).

The word-pairs were printed in random order in one column. They were presented in that order to half of the participants and in a reverse order to the other half.

Results

Mean predicted and actual response percentages are presented in Figure 1 for the zero, low, and high pairs. Predicted percentages were markedly inflated: Across all items, actual percentage of responding was 9.4%. In comparison, predicted percentage averaged 49.6%, fully 5 times as much! The difference was highly significant, $t(29) = 23.05$, $p < .0001$.

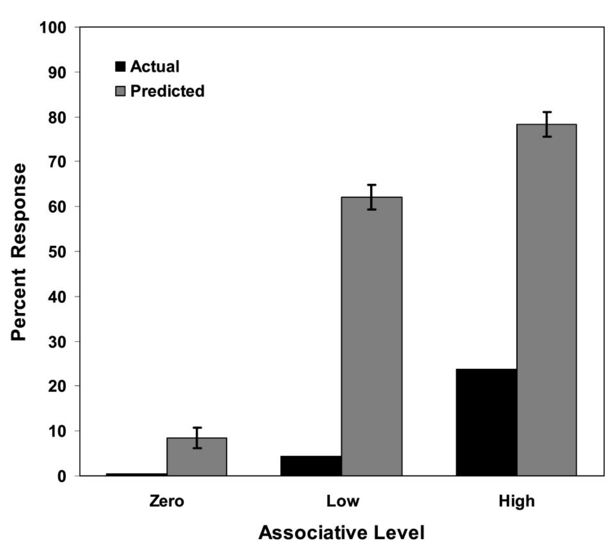

Figure 1. Mean actual and predicted response percentages for the zero-, low- and high- associative pairs. Error bars represent ±1 SEM (Experiment 1).

The prediction inflation effect was reliable across participants: For each and every participant, mean predicted percentage exceeded the mean actual percentage, $p < .0001$, by a binomial test. Predictions were inflated by a factor that ranged from 1:1.6 to 1:6.8 across participants. Prediction inflation was also consistently obtained across items: Mean predictions exceeded actual percentages for each and every one of the 45 items ($p < .0001$) by a binomial test. Predictions (for the low- and high-association pairs) were inflated by a factor that ranged from 1:2.4 to 1:19.5 across items.

As Figure 1 indicates, mean predictions increased with actual percentages, indicating that participants were sensitive to interitem differences (see below). Indeed, a one-way analysis of variance (ANOVA) comparing predictions for the zero, low, and high associative levels yielded, $F(2, 58) = 265.22$, $MSE = 151.48$, $p < .0001$. However, the steepest increase in prediction inflation was from the zero pairs to the low-association pairs. Thus, the difference between predicted and actual percentages amounted to 8.4%, 57.7%, and 54.6%, for the zero, low, and high pairs, respectively. Although inflation was significant even for the zero pairs, $t(29) = 3.66$, $p < .001$,[2] its magnitude exhibited a step-function in relation to actual rate of responding. It appears that a low relationship between the two members of a pair was sufficient to produce a very sizable inflation in predictions, an inflation that was no lower than that obtained for the high-association pairs. This pattern is similar to one that has been noted previously in connection with priming (Koriat, 1981) and judgments of learning (Koriat & Bjork, 2005, in press). Maki (2005a, 2005b) has also reported recently a very similar pattern of results for judgments of associative memory (JAM) that were obtained using several types of ratings on word pairs: JAM ratings, plotted as a function of normative cue−target probabilities were found to exhibit high intercepts, indicating inflated judgments, and shallow slopes, indicating relative insensitivity to differences in associative strength (see also Koriat, 1981).

In Figure 2 we plotted the results using a procedure similar to that of studies of the calibration of assessed probabilities (see Lichtenstein, Fischhoff, & Phillips, 1982). Mean over/underconfidence for each participant, computed as the weighted mean of the differences between the mean predicted percentage and the actual percentage for the 11 predicted percentage categories (0%−10%, 11%−20%, . . . 91%−99%, 100%; see Lichtenstein et al., 1982), averaged .40.

What about resolution, or discrimination accuracy (see T. O. Nelson & Dunlosky, 1991; Yaniv, Yates, & Smith, 1991), that is, the extent to which participants discriminate between items with different responding rates? Within-subject Pearson correlations, calculated across items, between estimated and actual percentages averaged .64 across participants, $t(29) = 25.15$, $p < .0001$. Except for 1 participant, for whom that correlation was .01, the remaining participants exhibited correlations that ranged from .47 to .81.

[1] Hebrew is written from right to left.

[2] It should be stressed that the analyses of prediction inflation is problematic in the case of the unrelated pairs because prediction could deviate only in one direction.

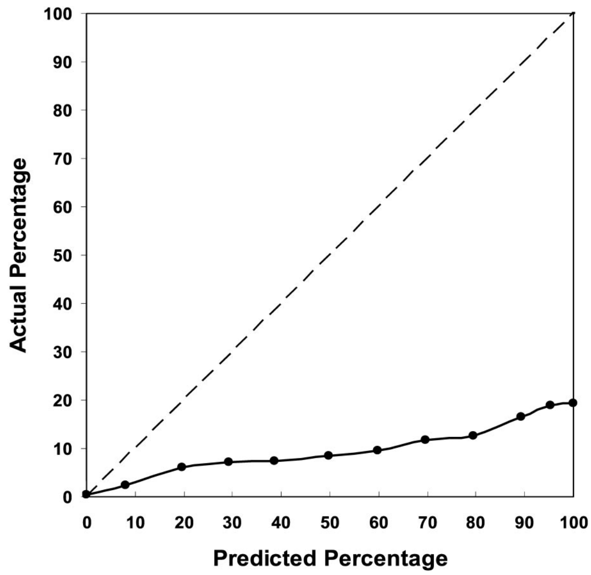

Figure 2. Calibration curve plotting actual response percentages as a function of mean predicted percentages (Experiment 1).

Thus, participants' predictions, although inflated, increased with increasing actual percentages, as also suggested by the results presented in Figures 1 and 2.

Discussion

The results of Experiment 1 documented a marked inflation of conditional predictions. This inflation was strongest for the low-association pairs, suggesting that even a weak connection between the stimulus and response words is sufficient to prime aspects of the stimulus that would make the response feel much more likely than it actually is. It should be stressed, however, that participants were not out of touch with reality because their inflated predictions were correlated to an appreciable degree with the actual probabilities of responding.

Experiment 2: Inclusion and Exclusion Judgments

Experiment 2 had two aims. The first was to eliminate a possible explanation of the inflated predictions observed in Experiment 1. That experiment included a set of pairs in which the responses were never observed in word association norms. The inclusion of such pairs might have resulted in a contrast effect that inflated predictions for the associated pairs. Thus, in Experiment 2 the zero-association pairs were removed from the list that was used.

The second aim of Experiment 2 was to examine the possibility that the prediction inflation bias would be reduced or eliminated when participants were induced to attend to alternative outcomes. Previous research has indicated that an effective strategy for debiasing judgments and reducing inflated predictions and overconfidence is to have participants "consider the opposite" (Hirt & Markman, 1995; Lord, Lepper, & Preston, 1984). Koriat, Lichtenstein, and Fischhoff (1980), for example, found that overconfidence in the correctness of one's chosen answers was reduced when people generated reasons contradicting these choices. Asking participants to produce counterarguments has been also found to reduce the explanation bias (e.g., Hoch, 1984; Koehler, 1991). Hirt and Markman (1995) found that considering any plausible alternative outcome for an event, not just the opposite outcome, helps in debiasing judgments.

Thus, a second condition was added in Experiment 2 in which participants were required to estimate the percentage of associative responses other than the focal, target response. This condition will be referred to as the *exclusion* condition, to distinguish it from the standard *inclusion* condition used in Experiment 1. If participants are able to detach themselves from the target word, and to consider seriously alternative responses, then the *complementary* values of these predictions (i.e., the difference between the observed percentage and 100%) should exhibit either no overestimation bias or even an underestimation of the occurrence of the focal target as a result of the perceived inflated likelihood of the alternative responses.

Method

Participants. Thirty-eight Hebrew-speaking participants took part in the experiment—24 high-school seniors and 14 first-year college students. They were tested during class meetings, and were assigned randomly to the two conditions, with 19 participants in each condition.

Procedure. The instructions for the inclusion condition were the same as those used in Experiment 1. Those for the exclusion condition were also the same except that participants were required to estimate what percentage of people "would not say the second word in response to the first word, but will say another word instead." As in Experiment 1, the word pairs were printed in random order in one column. They were presented in that order to half of the participants in each condition and in a reverse order to the other half.

Results

We first examine the estimates obtained in the inclusion condition. These will be referred to as *direct* estimates, to distinguish them from the complementary estimates based on the results of the exclusion condition (see Figure 3). Despite the elimination of the unrelated pairs, estimated percentages were still markedly inflated. Across all items, predicted percentages averaged 63.6% compared with 14.1% for actual percentages, $t(18) = 16.41$, $p < .0001$. As in Experiment 1, the magnitude of inflation was roughly the same for the low- and high-association pairs (50.8% and 48.2%, respectively). Also, as in Experiment 1, the inflation bias was very reliable. It was obtained for each of the participants ($p < .0001$) by a binomial test, and also for each of the 30 word pairs ($p < .0001$) by that test.

Did the elimination of the unrelated pairs reduce the overestimation effect? The inclusion condition of Experiment 2 yielded somewhat lower predictions than were found in Experiment 1: They averaged 55.3% for the low-association pairs and 71.9% for the high-association pairs (compared with 62.1% and 78.3%, respectively, in Experiment 1). An Experiment × Associative Level (low vs. high) ANOVA on these means, however, yielded only a near-significant effect for experiment, $F(1, 47) = 2.79$, $MSE = 364.81$, $p < .10$. Associative level yielded a significant effect, $F(1, 47) = 124.19$, $MSE = 53.00$, $p < .0001$, but the interaction was not significant ($F < 1$). Thus, the inflation bias observed in Experiment 1 was not due to the inclusion of unrelated pairs, and

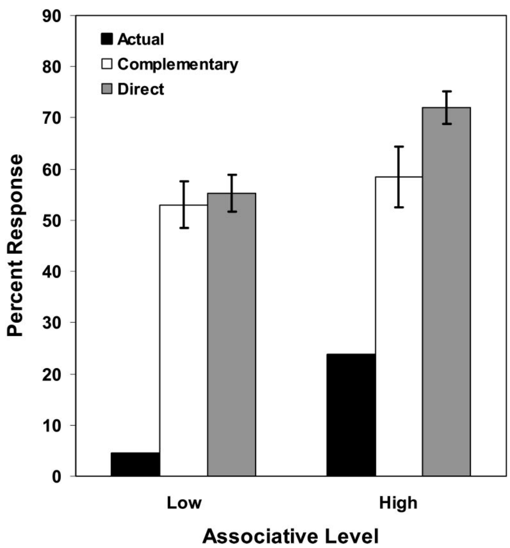

Figure 3. Mean predicted percentages for the direct and complementary conditions and mean actual response percentages for the low and high associative pairs. Error bars represent ±1 SEM (Experiment 2).

in any case, their inclusion did not seem to increase the degree of bias by much. The recent work of Maki (2005a) also indicates that JAM ratings are overestimated regardless of the inclusion of unrelated cue−target pairs in the list.

We turn next to the second aim of Experiment 2. Each of the predictions made by the exclusion group was subtracted from 100% to yield a complementary percentage score that could be compared with the direct predictions made by the inclusion group. The means of the complementary scores are also presented in Figure 3. The results clearly indicate that the overestimation bias was marked even for the exclusion condition: Participants underestimated the occurrence of alternative responses to the one stated. The complementary percentages averaged 55.7%, compared with 14.1% for the actual percentages, $t(18) = 8.68$, $p < .0001$. The effect was quite reliable: It was obtained for 17 out of 19 participants, with 1 participant yielding a tie ($p < .001$) by a binomial test, and also for each of the 30 pairs ($p < .0001$) by a binomial test. Thus, although participants were asked to focus on alternative responses to the focal outcome, the results suggest an inflation bias for that outcome.

Was the overestimation effect in Experiment 2 weaker for the exclusion than for the inclusion condition? A Condition (direct vs. complementary) × Associative Level (low vs. high) ANOVA yielded, $F(1, 36) = 1.93$, $MSE = 609.85$, *ns*, for condition, $F(1, 36) = 24.67$, $MSE = 93.30$, $p < .0001$, for associative level, and, $F(1, 36) = 6.48$, $MSE = 93.30$, $p < .05$, for the interaction. The interaction suggests that only for the high-association pairs was the tendency to overestimate the target response weaker for the exclusion than for the inclusion condition, $t(35) = 2.08$, $p < .05$. The low-association pairs, in contrast, yielded little difference between the two conditions, $t(35) = 0.40$.

Discussion

The prediction inflation effect was replicated in Experiment 2 even when the unrelated pairs were eliminated from the list of stimuli, so that the overestimated percentages observed in Experiment 1 cannot be attributed to a contrast effect resulting from the inclusion of the zero pairs. In fact, it might have been argued that the inclusion of such pairs (in Experiment 1) should suggest to participants that they have the option of providing very low estimates, thus resulting perhaps in lower estimates than would have been made otherwise. However, the results did not support that prediction either.

As far as the second aim of the experiment is concerned, the effects of the exclusion instructions were surprisingly negligible. These instructions helped reduce inflated estimates only for the high-association pairs and even then the reduction was quite small. This finding is at odds with the previously reported effects of manipulations that focus participants' attention on alternative outcomes (e.g., Hirt & Markman, 1995). One explanation for this discrepancy is that in those studies, specific alternatives were proposed, whereas in this experiment, participants were simply asked to think of other alternative outcomes. In terms of support theory, the hypothesis that is evaluated in the exclusion condition (producing a response other than the target response) is packed, unlike the hypothesis evaluated in the inclusion condition, which is explicit and specific. Unpacking an implicit hypothesis by making its components explicit has been found to increase its total probability markedly (Fiedler & Armbruster, 1994; Tversky & Koehler, 1994). Thus, perhaps if the implicit category "other than the target response" were to be split into several low-frequency components, an overestimation of that category would have been observed.

Another explanation, however, which is consistent with our thesis, is that the focal outcome modifies the construal of the condition, emphasizing those features of the condition that are consistent with it. Perhaps participants fail to detach themselves from the overriding assimilative influence of the target response. This may occur either because that response makes it difficult for them to access alternative responses or because it reduces the perceived plausibility of the accessed alternatives. Indeed, Sanna, Schwarz, and Stocker (2002) observed that when participants were instructed to produce many reasons why past events might have turned otherwise, they judged the alternative outcome as less likely than when they were instructed to produce only a few reasons. Presumably, the difficulty thinking about many alternative outcomes convinced participants that these outcomes are actually not very likely to occur (see also Wänke, Bless, & Biller, 1996). Perhaps in our experiment, too, the presentation of the target response made the accessibility of alternative responses more difficult, thus enhancing the judged likelihood of the target response.

We should stress, though, that it was particularly the low-association pairs that failed to yield a lower overestimation bias in the exclusion condition than in the inclusion condition. For these pairs, the actual percentage of the response was only 4.4, so it is not conceivable that participants were unable to access specific alternative responses to the cue words. Of course, it is possible that participants performed the exclusion task by judging the likelihood of the target outcome itself and then inferring the likelihood of making a different response. However, if they performed the task by considering alternative responses, then the inflated predictions for the low-association pairs would seem to suggest that participants also failed to appreciate the likelihood of the accessed responses. This interpretation is supported by the finding (see Experiment 4: Exploring the Backward Activation Account and a Potential Debiasing Procedure) that the inflation bias survived even when participants judged the likelihood of a target response after they themselves had produced a different response as their spontaneous first associate.

Experiment 3: The Overprediction Effect

If participants overestimate the likelihood of stated outcomes, then we should expect that the sum of the estimated probabilities of two alternative outcomes might sometimes exceed 1.00, an effect we label as an *overprediction effect.* Several previous studies have shown that probability judgments are often subadditive, summing to greater than the total possible probability (see Dougherty & Hunter, 2003; Mulford & Dawes, 1999). For example, Teigen (1974a, 1974b, 1983) showed that people commonly violate the convention that the probabilities assigned to an exhaustive set of mutually exclusive events should add up to 1. Only in the two-alternative case did a majority of participants give estimates that added up to unity (or 100%). As the number of alternatives increased, the total probability increased far beyond 100%, and the mean probability assigned to each alternative was almost independent of the number of alternatives. Sanbonmatsu et al. (1997) also showed that when different groups of participants estimated the number of votes that each of four candidates would receive, the average probability assigned to each candidate was .60 rather than .25.

Experiment 3A

Experiment 3A sought evidence for an overprediction effect by having participants estimate the percentage of occurrence of either one of two possible responses to the same cue word. For each of the cue words, the primary and secondary associates—that is, the two words that occurred as the most frequent responses according to word association norms—were used as target responses for two different groups of participants. We expected the total of the estimated percentages for some of the cue words to exceed 100%.

Method. Thirty-six Hebrew-speaking undergraduates (31 women and 5 men) participated in the experiment. The materials used included 20 stimulus words and their primary and secondary associates, compiled from Hebrew word association norms (Rubinsten et al., 2005). These were selected with the following constraints: (a) that the percentage of occurrence of the primary associate would not exceed 50% and (b) that the primary and secondary associates would not account for more than 70% of the responses. Examples of the items chosen (translated from Hebrew; percentage of occurrence in parentheses) are: copper − gold (28%), copper − iron (17%); chimney – smoke (49%), chimney – house (15%).

The 20 cue words were assigned randomly to two different lists such that each word was paired with its primary associate in one list, and with its secondary associate in the other list, and that each list included 10 primary response pairs and 10 secondary response pairs. In addition to these critical pairs, each list also included 10 low-association pairs and 10 unrelated pairs that were the same across the two lists.

The percentage of occurrence of the primary and secondary associates averaged 25.0 and 14.5, respectively across the 20 critical pairs. Percentage of occurrence averaged 4.5 for the low-association pairs (and 0 for the unrelated pairs).

Results. Considering first the results for the unrelated and low-association pairs, the estimated percentages for these pairs averaged 4.7% and 48.9%, respectively. The latter percentage was considerably higher than the actual percentage (4.5%), $t(35) = 16.48$, $p < .0001$.

Turning next to the critical items, estimated percentages averaged 73.5% and 69.4% for the primary and secondary associates, respectively, compared with the actual percentages of 25.0% and 14.5%. The estimated percentages exceeded the actual percentages for both the primary associates, $t(35) = 22.34$, $p < .0001$, and the secondary associates, $t(35) = 19.33$, $p < .0001$. Although the estimated percentage was significantly higher for the primary than for the secondary associates, $t(35) = 2.23$, $p < .05$, the difference between them was smaller than what might have been expected on the basis of the difference in their actual percentages.

We turn now to the primary aim of the experiment. The results clearly documented an overprediction effect: The sum of the estimated percentages of the primary and secondary associates averaged 142.9%, significantly higher than 100%, $t(35) = 9.03$, $p < .0001$. As can be seen in Table 1, the overprediction effect was observed for each of the 20 critical items. For these items, the summed predictions ranged from 114.1% to 170.8% (where the actual sums never exceeded 70.0%, and averaged 39.5% across items). An item-based t test comparing the estimated sums and actual sums of the primary and secondary associates yielded, $t(19) = 26.09$, $p < .0001$.

Note that the overprediction effect was obtained even though each of the stimulus words elicits many other responses according to word association norms. Table 1 lists (under "No. of associates") the number of different associative responses that are elicited by each stimulus word according to the norms. On average, each of the critical cue words elicits 26.2 responses, that is, 24 different responses in addition to the primary and secondary responses.

Experiment 3B

Although the overprediction effect was demonstrated in a between-subject design in Experiment 3A, we believe that it reflects a process that occurs within person (see also Teigen, 1974a, 1983). Several authors have argued against the usefulness of a within-participant design in the study of intuitive judgments (see Kahneman & Frederick, 2002), but we thought that it is important to examine whether the overprediction effect would be found even when each participant estimates the percentage of each of the alternative responses in turn. Experiment 3B, then, is a replication of Experiment 3A except that participants received both of the stimulus lists.

Method. Twenty Hebrew-speaking undergraduates (16 women and 4 men) were paid for participating in the experiment. The procedure was the same as in Experiment 3A except that both research forms were administered to the participants in counterbalanced orders, with a 3-min filler task between them.

Results. The estimated percentages averaged 66.2% and 62.6% for the primary and secondary associates, respectively. Their sum—128.8%—was significantly higher than 100%, $t(19) = 3.85$, $p < .001$. The overprediction effect was observed for each of the 20 critical items. An item-based t test comparing the estimated sums with the actual sums of the primary and secondary associates yielded, $t(19) = 20.36$, $p < .0001$.

The overprediction effect was observed for 16 out of the 20 participants ($p < .01$) by a binomial test. For the remaining

Table 1
Mean Actual and Estimated Percentage of Occurrence for the Primary and Secondary Associates, and for Their Sums, Listed by Item

Item	No. of associates	Primary associate		Secondary associate		Primary + secondary	
		Actual	Estimated	Actual	Estimated	Actual	Estimated
1	23	21.6	61.1	11.8	77.8	33.3	138.9
2	32	12.7	79.4	10.8	64.4	23.5	143.9
3	18	48.0	82.3	14.7	46.2	62.7	128.6
4	13	38.2	69.4	29.4	78.0	67.6	147.4
5	23	27.5	58.2	16.7	55.8	44.1	114.1
6	14	39.2	76.5	19.6	65.2	58.8	141.7
7	24	27.5	63.4	10.8	84.2	38.2	147.6
8	17	25.5	83.5	18.6	85.7	44.1	169.2
9	25	17.6	69.3	14.7	62.9	32.4	132.2
10	31	17.6	72.2	10.8	62.8	28.4	135.1
11	29	26.5	76.9	9.8	76.4	36.3	153.3
12	37	18.6	75.7	12.7	72.7	31.4	148.4
13	24	33.3	84.1	19.6	86.7	52.9	170.8
14	33	17.6	70.0	7.8	73.8	25.5	143.8
15	29	20.6	84.5	9.8	71.9	30.4	156.4
16	36	17.6	70.8	10.8	68.6	28.4	139.4
17	26	14.7	71.7	13.7	60.0	28.4	131.7
18	29	17.6	75.7	15.7	64.3	33.3	139.9
19	33	28.4	72.1	16.7	62.2	45.1	134.3
20	27	29.4	73.7	15.7	68.1	45.1	141.7
Mean	26.2	25.0	73.5	14.5	69.4	39.5	142.9

Note. No. of associates = the total number of associates elicited by the stimulus word of each item (Experiment 3A).

participants the summed predictions were 66.3%, 83.7%, 84.1%, and 86.0%. The magnitude of overprediction was reliable: The correlation across participants between the mean estimates made for the critical items on the first and second forms was .65 ($p <$ .005), suggesting that participants differed reliably in the tendency to provide high or low estimates or, perhaps, in the ability to construe the condition−outcome link effectively.

Prediction inflation, however, was weaker in the within-participant design: An item-based analysis comparing the summed predictions in Experiments 3A and 3B yielded, $t(38) = 3.30$, $p <$ .005. This pattern came about because the predicted percentages in Experiment 3B decreased from the first list (136.8%) to the second list (120.7%), $t(19) = 3.40$, $p < .005$, in an item-based analysis. It is possible that participants attempt to avoid overprediction when it becomes more transparent (see also Bastardi & Shafir, 1998, for an analogous finding).

Discussion

The results of Experiment 3A indicate that participants overestimated the likelihood of each of the alternative responses to the cue word to the extent that their estimated percentages of occurrence totaled more than 100%. These results suggest that alternative outcomes are not perceived to compete with each other in a zero-sum fashion. Experiment 3B replicated this pattern in a within-person design. The results indicate that having assigned very high estimates to one response did not prevent participants from assigning very high estimates to an alternative response to the same cue.

The failure to perceive alternative outcomes as being in complete competition may derive from a process similar to that underlying the encoding specificity principle (Tulving & Thomson, 1973): It is as if the cue word *chimney*, for example, is encoded differently when paired with *smoke* as a potential response than when paired with *house* as a potential response. This happens, in our view, because different target words bring to the fore aspects of the cue word that are consistent with it.

Experiment 4: Exploring the Backward Activation Account and a Potential Debiasing Procedure

Experiment 4 had two aims. The first was to obtain evidence for the hypothesis that prediction inflation results in part from a backward activation in which the outcome brings to the fore aspects of the condition that are less likely to come forward when that outcome is not stated (see Koriat & Bjork, 2005). To test this idea, we used word pairs with an asymmetric association such that the association from Word A to Word B is much stronger than the association from B to A. For example, the words *cheddar* and *cheese* are asymmetrically related in that 92% of the people who are presented with the word *cheddar* respond with *cheese*, whereas only 5% of those who are presented with the word *cheese* respond with *cheddar* (D. L. Nelson, McEvoy, & Schreiber, 1999). We expect prediction inflation to be particularly strong for backward-ordered pairs (cheese – cheddar) because the presence of the response (cheddar) is assumed to activate those aspects of the stimulus (cheese) that are less likely to dominate when that stimulus appears alone.

The second aim of Experiment 4 was to explore the effectiveness of a debiasing procedure—requiring participants to generate their own responses. A generation condition was included, in which participants gave their first association to the cue word before they were presented with the stimulus−response pair and asked to estimate the percentage of people who would respond with the stated target word.

The expectations about the effects of the generation manipulation are not clear. On the one hand, if the participant's own response differs from the focal, target response, we might expect the overestimation bias to disappear or even change to underestimation. This expectation is consistent with findings indicating that inducing participants to consider alternative outcomes can reduce inflated probabilities (see Hirt & Markman, 1995). As noted earlier, Tversky and Koehler (1994) argued that overconfidence in the correctness of one's judgment sometimes arises because the target hypothesis is specified, whereas its alternatives are not. However, in the generation condition, the production of a response that does not match the target response may be considered to constitute a specification of an alternative hypothesis (unlike the exclusion condition of Experiment 2) and would therefore be expected to attenuate the probability of the target hypothesis.

On the other hand, our conceptual framework leads us to doubt the effectiveness of the generation manipulation in eliminating prediction inflation. Even when a participant's own response differs from the focal response, the presentation of the focal response in conjunction with the cue word may preempt the experience gained from generating one's own response: Having produced *house* in response to *chimney* need not make *smoke* less likely as a response to *chimney* when the pair *chimney* − *smoke* is presented. Thus, it is of theoretical interest to see whether the prediction inflation bias survives the effects of a generation manipulation.

Method

Participants. Forty University of Haifa students (24 women and 16 men) participated in the experiment—11 for course credit and 29 as paid participants. They were assigned randomly to the generation and control conditions, with 20 participants in each group.

Materials. A list of 90 Hebrew word pairs was compiled, consisting of 30 unrelated word pairs and 60 asymmetrically associated pairs. The asymmetrical pairs were chosen on the basis of the norms collected for Experiment 3 of Koriat and Bjork (in press). For these pairs, the percentage responding according to the norms was 59.9% in the dominant direction and only 2.0% in the opposite direction. The pairs were divided into two sets of 30 pairs each. One set was presented in the forward direction (i.e., such that the strongest association was from the cue word to the target word) and the other set was presented in the backward condition (with the strongest association being from the target to the cue), with the assignment of the two sets to the forward and backward conditions counterbalanced across participants. In addition, all participants received the 30 unrelated pairs.

Apparatus and procedure. The experiment was conducted on a personal computer. The experimental materials were displayed on a computer screen. The instructions for the control condition were similar to those of Experiment 1, except that participants spoke their estimates into a microphone, and the next pair appeared 1 s after the experimenter entered the response. The procedure for the generation group was the same except that each trial began with a generation phase: Only the cue member of the pair was presented, and participants were asked to speak into the microphone the first word that came to mind in response to the cue word. Once the experimenter had recorded the response, the stimulus pair disappeared and was replaced by the stimulus−response pair for the estimation task. The order of presentation of the items was random except that the same order was used for each pair of yoked participants belonging to the two different conditions.

Results

Figure 4 presents mean actual and predicted percentages of occurrence for the generation and control conditions. Two questions will be addressed. First, did control participants exhibit a stronger inflation bias for the backward pairs than for the forward pairs? Second, did the generation manipulation help in eliminating this bias, and particularly so for the backward pairs?

Prediction inflation for the forward and backward pairs. With regard to the first question, the results of the control condition indicated an inflation bias for each of the associative classes, $t(19) = 3.97$, $p < .001$, for the unrelated pairs, $t(19) = 15.35$, $p < .0001$, for the backward pairs, and, $t(19) = 11.24$, $p < .0001$, for the forward pairs. Consistent with our expectations, the backward pairs yielded a very strong inflation: The predicted percentage averaged 64.0% when the actual percentage was only 2.0%. The inflation was less severe for the forward pairs, for which the respective values were 81.1% and 59.9%. Because no error variance is available for the actual percentages, we performed the analyses on the estimated−actual differences calculated for each participant. A comparison of the difference scores for forward and backward pairs yielded, $t(19) = 13.61$, $p < .0001$, indicating a stronger bias for the backward pairs.

The inordinately high predictions observed for the backward pairs might have been seen to derive from a simple statistical regression in which small frequencies tend to be overestimated (Fiedler & Armbruster, 1994): Because the actual percentages for the backward pairs were very low, a noisy judgment distribution would be more likely to yield greater inflation for these pairs than for the forward pairs. The results for the unrelated pairs, however, argue against this interpretation: The predictions for these pairs averaged 8.1% (when the actual percentage was zero), much lower than what was found for the backward pairs (64.0%).

The effects of response generation. Turning next to the second question: Did the generation task eliminate the prediction inflation bias? Clearly, this was not the case (see Figure 4). A sizable

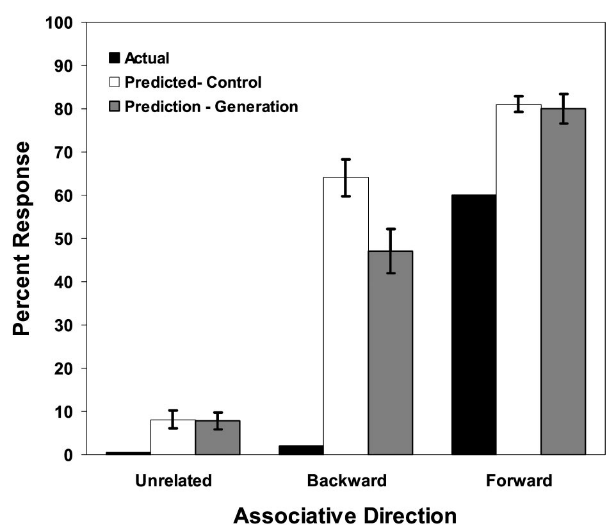

Figure 4. Mean actual and predicted response percentages for the control and generation conditions for the unrelated, backward and forward pairs. Error bars represent ±1 SEM (Experiment 4).

inflation was obtained for the forward and backward pairs, $t(19) = 6.09$, $p < .0001$, and, $t(19) = 9.07$, $p < .0001$, respectively. Furthermore, as was the case for the control participants, the bias was stronger for the backward pairs (45.1%) than for the forward pairs (20.1%), A t test comparing the predicted−actual differences yielded, $t(19) = 7.81$, $p < .0001$.

However, the generation manipulation did reduce somewhat the magnitude of inflation, but only for the backward pairs: The reduction amounted to 0.4%, 1.1%, and 17.0%, for the unrelated, forward, and backward pairs respectively, and was significant only for the backward pairs, $t(38) = 2.61$, $p < .05$. However, the prediction inflation demonstrated by the generation participants for the backward pairs is dramatic and is especially striking when one considers that the word generated by participants in that condition, prior to their predicting the behavior of others, was virtually always a different word, as we discuss below.

Comparing same and different trials across all items. To clarify the effects of the generation task, we distinguished between trials in which the response generated matched the target word (same), and those in which it differed from it (different). The proportion of same responses across all items averaged .25 across participants. On average, participants made lower predictions when their own response differed from the target response than when it matched it (see Figure 5). However, same responses were more likely to occur for pairs for which the actual, normative percentage was high, so that the corresponding actual percentages also differed, as shown in Figure 5. A two-way ANOVA on these means,[3] Response Agreement (same vs. different) × Measure (predicted vs. actual) yielded significant effects for response agreement, $F(1, 19) = 749.03$, $MSE = 71.76$, $p < .0001$, and for measure, $F(1, 19) = 69.22$, $MSE = 179.64$, $p < .0001$, but the interaction was significant only at the .20 level, $F(1, 19) = 1.83$, $MSE = 26.17$. A comparison of the predicted and actual percentages indicated that the overestimation bias was significant for both same trials, $t(19) = 7.97$, $p < .0001$, and different trials, $t(19) = 7.57$, $p < .0001$.

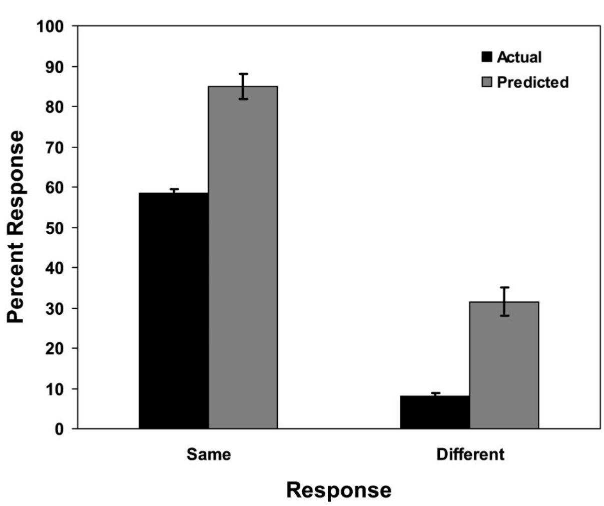

Figure 5. Mean actual and predicted response percentages for Same and Different responses. Error bars represent ± SEM (Experiment 4).

Comparing same and different trials for forward and backward pairs. We shall now examine these results in greater detail, focusing on the comparison between the generation and control conditions for the forward and backward pairs.

As would be expected, the likelihood of generating the target word was much smaller for the backward than for the forward pairs, .048 versus .687, respectively. (The respective normative probabilities were .020 and .599.) Whereas all 20 participants produced same responses for some of the forward pairs, only 15 participants did so for some of the backward pairs. This difference alone, however, cannot account for the stronger effect of generation on the backward than on the forward pairs, as suggested by the following analysis. In that analysis, we used only the 15 participants for whom there were same responses for some of the backward pairs. Each of these was yoked randomly with one control participant such that all items for the control participant were also classified as same or different according to the classification of these items for the yoked (generation) participant. Mean predictions for the control and generation conditions are plotted in Figure 6 for forward and backward items classified as same and different.

An Associative Direction (forward vs. backward) × Response Agreement (same−different) × Condition (control vs. generation; treated also as a repeated factor) ANOVA yielded, $F(1, 14) = 20.07$, $MSE = 321.29$, $p < .001$, for associative direction, with forward pairs yielding higher overall predictions (78.3%) than backward pairs (63.7%). There was also a significant effect for response agreement, $F(1, 14) = 31.37$, $MSE = 146.36$, $p < .0001$, which interacted with condition, $F(1, 14) = 52.00$, $MSE = 84.64$, $p < .0001$: The production of a response that matched the target exacerbated the inflation bias by 5.1%, whereas the generation of a different response reduced predictions by 19.1%. The triple interaction only approached significance, $F(1, 14) = 2.66$, $MSE = 155.72$, $p < .13$, perhaps reflecting the observation that only the reduced prediction for different responses was significant, $t(14) = 3.12$, $p < .01$, whereas the increased prediction for same responses was not, $t(14) = 0.26$. Thus, the generation manipulation tended to exacerbate the inflation bias for same responses, but not significantly so. When participants produced a different response, however, this was equally effective in reducing overestimation for both the forward and backward pairs.

We compared next the forward and backward pairs in the magnitude of prediction inflation when the production of same or different response was controlled. For same trials, a Measure (predicted vs. actual) × Associative Direction (forward vs. backward) ANOVA (based on 15 participants) yielded a stronger inflation for the backward than for the forward pairs. For the forward pairs, predicted and actual percentages averaged 84.6% and 62.3%, respectively, $t(14) = 6.44$, $p < .0001$. The respective means for the backward pairs were 74.8% and 4.3%, respectively, $t(14) = 11.95$, $p < .0001$. A Measure × Associative Direction ANOVA yielded a significant interaction, $F(1, 14) = 161.77$, $MSE = 54.07$, $p < .0001$. Prediction inflation was about three times larger for backward pairs (amounting to 70.6%) than for the

[3] We could perform ANOVA in this case because there was error variance for the actual percentages.

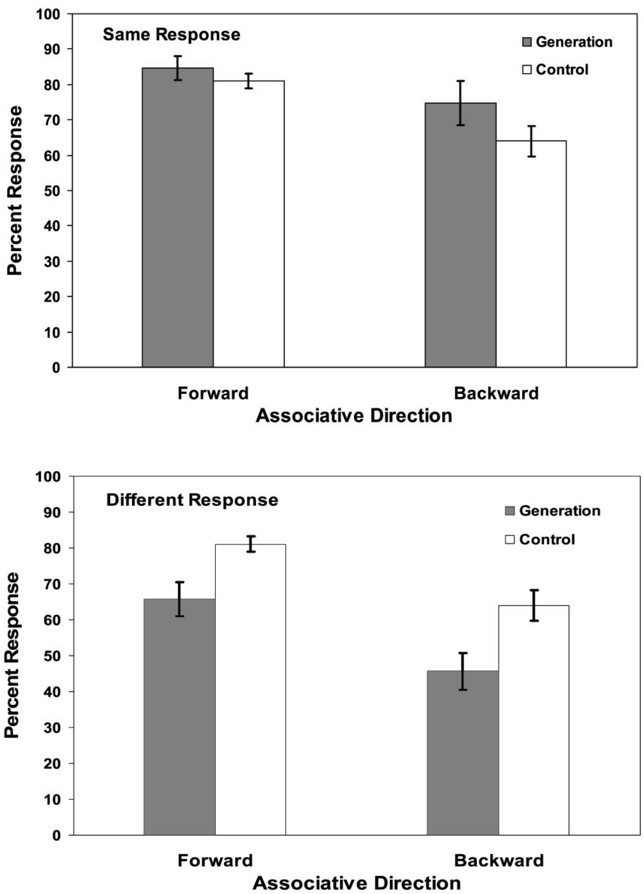

Figure 6. Mean predicted response percentages for the forward and backward pairs in the generation and control conditions. The results are plotted separately for Same responses (top panel) and Different responses (bottom panel). Error bars represent ± 1 SEM (Experiment 4). (See text for a description of how Same and Different responses were defined for the control condition.)

forward pairs (22.3%). A *t* test on the difference between estimated and actual percentages yielded, $t(14) = 12.72$, $p < .0001$, for the forward−backward contrast.

A similar comparison for different trials (based on all 20 participants) also indicated a stronger inflation prediction for the backward pairs. A Measure × Associative Direction ANOVA on these trials, yielded significant effects for measure, $F(1, 19) = 38.43$, $MSE = 403.57$, $p < .0001$, and for associative direction, $F(1, 19) = 627.27$, $MSE = 41.53$, $p < .0001$, but the interaction was also highly significant, $F(1, 19) = 175.96$, $MSE = 28.89$, $p < .0001$. Prediction inflation was almost 4 times larger for the backward pairs (43.8%) than for the forward pairs (11.9%). A *t* test comparing the difference between estimated and actual percentages yielded, $t(19) = 13.26$, $p < .0001$, for the forward−backward contrast.

In sum, the backward pairs produced a more severe inflation than the forward pairs even in the generation condition, and also when the production of same or different responses was controlled. These results are consistent with the idea that the inflation is due in part to a process that occurs ad hoc as a result of activations emanating from the target outcome. Thus, after producing one's own response, the presentation of the cue along with the same or a different response may act to prime and reveal aspects of the cue that are less likely to affect one's (or others') actual responses.

This interpretation is reinforced by the results for the unrelated pairs: Mean predicted and actual percentages for these pairs were 7.8% and 0%, respectively. Note that there were no same responses for the unrelated pairs, so that in this respect they are similar to the different trials of the backward pairs. However the respective means for the latter pairs were 45.7% and 1.9%. Thus, the availability of a backward association results in strongly inflated predictions despite the experience of having produced a different response than the target response. These results also argue against an interpretation of the inflated predictions for the backward pairs in terms of statistical regression.

Discussion

The first aim of Experiment 4 was to gain insight into the mechanism underlying prediction inflation. The results were consistent with idea that this inflation derives from a posteriori associations that are activated by the target response. In the case of backward-associated pairs, these associations tend to be inordinately strong in comparison with the a priori associations that are activated by the cue when it appears alone. For example, the word *cheese* evokes such associates as *mouse*, *cracker*, and *milk*, more often than it does *cheddar*. However, when the pair *cheese* − *cheddar* is presented, *cheddar* tends to activate backward associations that give rise to the feeling that it is a very likely response. Indeed, in the control condition of Experiment 4, the estimated percentage was inflated by a factor of 1.3 for the forward pairs and by a factor of 31.9 for the backward pairs.

The second aim of Experiment 4 was to evaluate the effectiveness of the generation manipulation in alleviating prediction inflation. This manipulation can be assumed to give participants first-hand experience with the task whose outcome they are subsequently asked to predict, making them aware of the likelihood of responses other than the target response. For example, participants in the generation condition produced a word that matched the target word in less than 5% of the trials for the backward pairs, and this may be expected to make participants aware of the low probability of the target response.

The generation of one's own response, however, was barely effective in reducing prediction inflation, and the overestimation bias remained high even when the participant's response differed from the target word. Maki (2005b) also observed that asking participants to rate the likelihood of a word association response in the presence of other possible responses did not eliminate the overestimation of the occurrence of that response. Also, in their correlational analysis of a large database, D. L. Nelson, Dyrdal, and Goodmon (2005) obtained results suggesting that although forward free-association probabilities seem to be affected by the strength of competing associates, subjective ratings of the associative similarity between the words were not. Thus, conditional predictions seem to behave more like similarity ratings, being relatively indifferent to the presence of competing associates.

These observations are consistent with the failure of the exclusion condition to eliminate the prediction inflation for the focal target (Experiment 2). Taken together, the results of Experiments

2 and 4 suggest that the presentation of the target response along with the cue word largely preempts the experience gained from the consideration of other potential responses. Consistent with this interpretation, a stronger inflation was observed for the backward pairs than for the forward pairs even when the participant's generated response differed from the target response.

Experiment 5: Generating Two Associates

Why did participants fail to apply their experience in generating a response when they predicted the likelihood of the target response? One possibility is that the generation of a single response is not sufficient to make participants aware of the variety of responses available, a variety that was clearly noticeable across participants. In the nomenclature of support theory (Tversky & Koehler, 1994), generating two (different) responses is more effective for unpacking the alternative hypothesis than generating one response. Indeed, the results of several previous studies suggest that the perceived likelihood of an outcome can be reduced by simulating causal scenarios to a variety of alternative outcomes as against simulating several scenarios to the same outcome (Dougherty, Gettys, & Thomas, 1997; Hirt & Markman, 1995; Levy & Pryor, 1987). In line with these and other results (Dougherty & Hunter, 2003; Hirt & Markman, 1995) we examined whether asking participants to produce two associations to the stimulus word might be effective in reducing inflation prediction. It is interesting to see, then, whether the prediction inflation bias survives even when both of the participant's responses differ from the target response.

Method

Participants. Twenty University of Haifa students (13 women and 7 men) participated in the experiment, 16 for course credit, and 4 were paid for their participation.

Procedure. The procedure was the same as that of the generation condition of Experiment 4 except that participants were asked to provide two associations instead of one.

Results

Participants provided only a single association in 1.0% of the trials, and these trials were eliminated from the analyses. Figure 7 presents mean predicted and actual percentage of occurrence. Included in this figure are also the results for the control condition of Experiment 4. To our surprise, the results of Experiment 5 were similar to those of Experiment 4, yielding marked prediction inflation. Estimated percentages averaged 45.7% overall in Experiment 5 in comparison with 45.0% in Experiment 4 (the respective actual percentage was 20.7%). Generating two associations rather than one increased somewhat the estimates for the forward pairs (from 80.0% to 87.3%) but decreased the estimates for the backward pairs (from 47.1% to 44.1%). Neither of these effects, however, was significant, $t(38) = 1.88$, $p < .07$, and, $t(38) = 0.46$, *ns*, respectively.

As would be expected, participants' responses in Experiment 5 matched the target response more often than in Experiment 4: The proportion of trials in which one of the two generated responses matched the target averaged .002, .083, and .805 for the unrelated, backward, and forward pairs, respectively (the respective percentages in Experiment 4 were .000, .048, and .687).

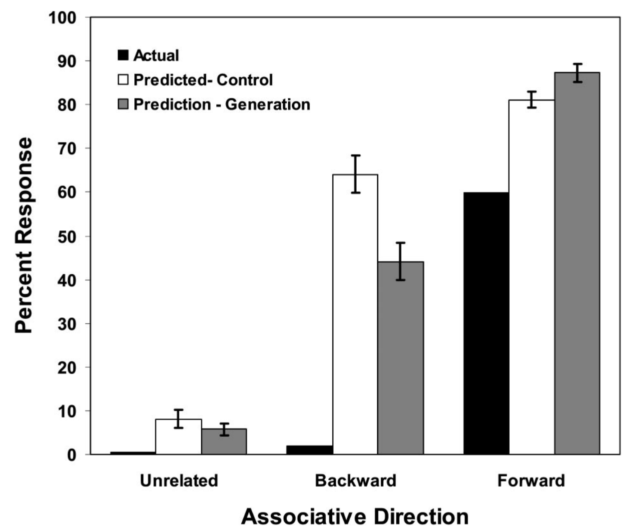

Figure 7. Mean actual and predicted response percentages plotted separately for the control condition of Experiment 4 and the generation condition of Experiment 5 for the unrelated, backward, and forward pairs. Error bars represent ±1 SEM.

As in Experiment 4, participants made higher predictions when their response matched the target response than when it differed from it. For the forward pairs predicted percentages for same and different responses averaged 91.7% and 70.5%, respectively, when the respective actual percentages were 61.8% and 51.5%. A two-way ANOVA, Response Agreement (same vs. different) × Measure (predicted vs. actual) yielded, $F(1, 19) = 33.97$, $MSE = 17.45$, $p < .0001$, for the interaction, indicating a stronger inflation for same than for different responses. A similar pattern was obtained for the backward pairs (using only 16 participants who produced same responses for some of the pairs): Mean predicted percentages for the same and different responses were 78.9% and 46.3%, compared with 4.1% and 2.0%, respectively, for the actual percentages, $F(1, 15) = 40.40$, $MSE = 91.96$, $p < .0001$, for the interaction.

Nevertheless, even the different trials yielded inflated predictions. For the forward pairs, predicted percentages averaged 70.5%, when the actual percentages averaged 51.5%, $t(19) = 6.97$, $p < .0001$. The respective means for the backward pairs were 41.5% and 1.8%, $t(19) = 9.96$, $p < .0001$. Furthermore, as in Experiment 4, the forward−backward difference was also replicated: Prediction inflation amounted to 19.0% for the forward pairs and to 39.7% for the backward pairs. A t test on the difference between estimated and actual percentages yielded, $t(19) = 5.66$, $p < .0001$, for the forward−backward contrast.

To clarify the effects of the generation task, we compared the results of Experiment 5 with those of the control group of Experiment 4 using a similar yoking procedure as that used in that experiment. There were only 16 participants in Experiment 5 who produced same responses for some of the backward pairs, and each was yoked randomly with one participant in the control condition of Experiment 4. An Associative Direction (forward vs. backward) × Response Agreement (same−different) × Condition (control vs. generation; treated also as a repeated factor) ANOVA, yielded similar results to those of Experiment 4. There was a significant effect for associative direction, $F(1, 15) = 51.10$, $MSE = 150.72$, $p < .0001$, with forward pairs yielding higher overall predictions (80.7%) than backward pairs (65.2%). There was also a significant effect for response agreement, $F(1, 15) = 67.38$, $MSE = 103.27$, $p < .0001$, which interacted with condition, $F(1, 15) = 42.62$, $MSE = 110.44$, $p < .0001$. The production of a response that matched the target exacerbated the inflation bias by 9.5%, whereas the generation of a different response reduced predictions by 14.8%. In addition, the Response Agreement × Associative Direction interaction was also significant, $F(1, 15) = 7.17$, $MSE = 51.00$, $p < .05$, suggesting a larger backward−forward difference for different responses (18.9%) than for same responses (12.1%). The triple interaction only approached significance, $F(1, 15) = 2.56$, $MSE = 67.81$, $p < .14$. Unlike in Experiment 4, here same responses increased predictions significantly, $t(15) = 2.61$, $p < .05$, and different responses reduced predictions significantly, $t(15) = 2.73$, $p < .05$.

Nevertheless, as in Experiment 4, predictions were markedly inflated even for different responses, and they were considerably more inflated for the backward pairs than for the forward pairs. Thus, focusing on different responses, predicted and actual percentages averaged 67.3% and 45.4%, respectively, for forward pairs, and 46.2% and 5.8%, respectively, for backward pairs. For same responses predicted and actual percentages averaged 91.5% and 59.1%, respectively, for forward pairs, and 78.1% and 5.7%, respectively, for backward pairs.

Discussion

The results of Experiment 5 replicated very closely those of the generation condition of Experiment 4. The requirement to generate two associations reduced the estimates but only when none of the produced associations matched the target word. The results also replicated the two main findings from the control condition of Experiment 4: First, predictions were markedly inflated even when participants produced two responses neither of which matched the target response. Second, the magnitude of this inflation was considerably larger for the backward than for the forward pairs.

It is noteworthy that the overestimation bias was no weaker in Experiment 5 than in Experiment 4, in which participants produced a single association. This was true even when the comparison was confined to different trials: A Condition (one association vs. two associations) × Measure (predicted vs. actual) ANOVA for different trials yielded, $F(1, 38) = 137.17$, $MSE = 72.29$, $p < .0001$, for measure, $F(1, 38) = 2.38$, $MSE = 121.57$, *ns*, for condition, and, $F < 1$, for the interaction. Prediction inflation amounted to 21.2% for the two-association condition (Experiment 5) and to 23.4% for the single-association condition (Experiment 4).

Experiment 6: Generating Two Associates in the Presence of the Target Response

In discussing the results of Experiments 4, we proposed that people fail to benefit fully from their own experience when generating a prediction about others because when the target response is later presented, it activates aspects of the cue word that are consistent with it, leading to the feeling that it too constitutes a highly probable response. How then can people be induced to avoid the influence of the backward-activation process and bring their experience to bear on their judgments?

One final generation procedure was explored in Experiment 6. Participants were shown the cue−target pair and asked to produce two additional alternative responses to the same cue word. Thus, the generation of these responses took place in the presence of the entire cue−target pair. Perhaps under these conditions participants can escape the process in which the presentation of the target brings to the fore new aspects of the cue word that were not realized before. Thus, Experiment 6 was similar to Experiment 5 except that participants gave two associations to the cue word in the presence of the target whose likelihood they were later asked to judge.

Method

Twenty University of Haifa students (14 women and 6 men) participated in the experiment for course credit. The procedure was the same as that of Experiment 5 except that participants first saw the entire cue−target pair, and were instructed to say the first two words that came to mind in response to the cue word other than the one presented.

Results

There were 26 trials (1.4%) in which participants failed to give two different associations. The results from these trials were eliminated from the analyses.

Figure 8 presents mean predicted and actual percentage of occurrence. Included in this figure are also the results from Experiment 5. The results of Experiment 6 were very similar to those of Experiment 5, demonstrating a strong prediction inflation. For the forward pairs, predicted percentages averaged 82.6%, compared with a mean actual percentage of 59.9%, $t(19) = 10.55$, $p < .0001$. The respective means for the backward pairs were 55.8% and 2.0%, $t(19) = 17.33$, $p < .0001$. The forward−backward difference was also replicated: Prediction inflation was more se-

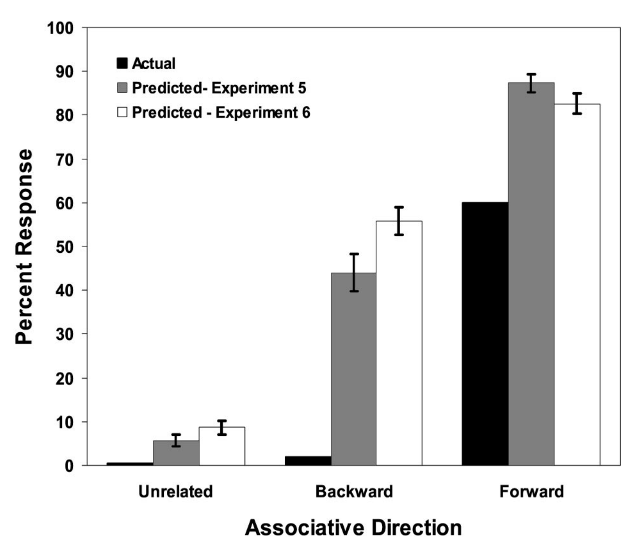

Figure 8. Mean actual and predicted response percentages for Experiment 5 and Experiment 6 for the unrelated, backward, and forward pairs. Error bars represent ±1 SEM.

vere for the backward pairs (53.8%) than for the forward pairs (22.6%), $t(19) = 11.90$, $p < .0001$, for the estimated−actual difference.

Overall, across all pairs, mean estimated percentages were very similar for Experiments 5 and 6: 46.0% and 48.9%, respectively, $t(38) = 1.09$, *ns*. Asking participants to generate associations in the presence of the entire pair rather than in the presence of the cue alone decreased somewhat the estimates for the forward pairs (from 87.3% to 82.6%), $t(38) = 1.55$, $p = .13$, but increased the estimates for the backward pairs (from 44.0% to 55.8%), $t(38) = 2.27$, $p < .05$. However, it should be noted that whereas in Experiment 5 the participants' responses could be the same as the target response or different, in Experiment 6 participants were explicitly instructed to give two associations that differed from the target association. Thus, Experiment 6 yielded, in fact, a stronger inflation bias (28.4%) than the different trials in Experiment 5 (21.3%). However, this difference likely stems from the fact that participants in Experiment 5 gave different responses to a selective set of items, for which the actual percentage of responding averaged 5.5%, whereas in Experiment 6 they gave different responses to all items, so that the actual percentage of responding with the target was much higher: 20.7%.

In fact, in comparing the results of Experiment 6 with those of the control group of Experiment 4, it would seem that the generation task used in Experiment 6 had little effect in terms of alleviating the inflation bias. Across the backward and forward pairs, the estimates in Experiment 6 averaged 69.2%, compared with 72.6% for the control condition of Experiment 4, $t(38) = 0.94$, *ns*. In comparison with that condition, the generation task of Experiment 6 inflated predictions slightly (by 1.5%) for the forward pairs but reduced predictions (by 8.2%) for the backward pairs.

Discussion

By and large, the elicitation of one's own responses in the presence of the focal outcome failed to yield a greater alleviation of prediction inflation than when the focal outcome was revealed only after the generation task. A possible explanation for this failure is that the presence of the target constrains the generation of "different" responses so that these responses tend to be sampled from the semantic space defined by the cue−target association. A cursory examination of the generated associates suggests that such might have been the case. If so, this would imply that somewhat different mechanisms are responsible for the relative ineffectiveness of the generation task in Experiments 5 and 6. Whereas in Experiment 5 participants failed to apply the experience from the generation task to the prediction task, in Experiment 6 they failed to exploit the generation task in full. It is also possible that because participants were exposed to the target before they produced their own responses, they experienced greater difficulty in producing their own responses than in Experiment 5, and this may have also contributed to increasing the perceived likelihood of the target response (see Sanna et al., 2002). Needless to say, these speculations require further study. What is important to note is that the generation tasks used in Experiments 4, 5, and 6 were by and large ineffective in eliminating or reducing markedly the prediction-inflation effect.

General Discussion

The results reported in this article support and reinforce previously reported findings indicating that people generally overestimate the likelihood of occurrence of the events that they are asked to judge. One advantage of the experimental paradigm we used in this study derives from the availability of a criterion measure for a large number of condition−outcome combinations against which participants' estimates could be compared. The comparison indicated a pervasive and marked tendency to overestimate the occurrence of the judged target outcomes (see also Maki, 2005a, 2005b).

In what follows we first discuss some of the previous ideas regarding the overestimation of the probability of future outcomes. We then focus on conditional predictions, discussing our proposal that one mechanism that contributes to the inflation of future probabilities in this case derives from backward activations. Finally, the results of the attempts to reduce prediction inflation are discussed in terms of how they bear on the processes underlying this inflation.

Overestimating the Likelihood of Target Outcomes

There has been ample evidence for the tendency of people to overestimate the likelihood of stated events. Much of that evidence comes from studies that examined the effects of imagination or explanation on the subjective probability of the judged events. The overestimation observed has been explained in terms of the selective testing and confirmation of a hypothesis (Fischhoff & Beyth-Marom, 1983; Hirt & Markman, 1995; Koriat et al., 1980; Sanbonmatsu et al., 1997). It has been proposed that the imagination of a future event and its explanation lead people to build a causal scenario that makes the event feel more likely than it actually is (Koehler, 1991). Indeed, results suggest that people test hypotheses by gathering supportive evidence and neglecting evidence for alternative hypotheses (e.g., Hoch, 1985; Sanbonmatsu et al., 1997; Wilson & LaFleur, 1995).

Other accounts invoked the operation of the availability heuristic (Tversky & Kahneman, 1973): Imagining or explaining an outcome is said to increase its accessibility. The ease of recalling or reconstructing that outcome enhances, then, its judged likelihood. Indeed, Sherman, Cialdini, Schwartzman, and Reynolds (1985) found that participants who were asked to imagine contracting a disease made higher likelihood judgments when the symptoms were easier to imagine than when they were more difficult to imagine.

However, people seem to overestimate the likelihood of future events even when they are not explicitly instructed to explain these events. Koehler (1991), for example, argued that when people are asked to evaluate a hypothesis, they typically act as if that hypothesis were true; it is the adoption of such a conditional reference frame, rather than the construction of an explanation per se, that causes increased confidence in the target hypothesis. Fiedler (2000) also argued that the mere considering of a particular proposition induces the belief that the proposition is true. Consistent with these suggestions, our results clearly showed that participants' predictions were consistently inflated in the absence of instructions that solicited explanations of the judged outcomes. Furthermore, these predictions continued to be inflated even when participants' own responses suggested different outcomes from those that had to be judged. This latter effect, however, might be specific to conditional predictions, as we now discuss.

Conditional Predictions and Their Inflation

In this study we focused on conditional predictions. Such predictions are very common in everyday life, particularly when people anticipate some future developments and must assess the likelihood of certain outcomes given these developments. Previous discussions of prediction inflation have focused on the changes that occur in the perception of an outcome as a result of explicitly mentioning that outcome. For example, Tversky and Koehler (1994) proposed that "different descriptions of the same event may call attention to different aspects of the outcome and thereby affect their relative salience" (p. 548). Thus, according to their support theory, unpacking a category (e.g., death from unnatural causes) into its components (e.g., homicide, fatal car accident) reminds people of possibilities that they might not have considered otherwise.

In the present study, in contrast, we focused on the changes that may occur in the perception of the condition as a potential source of prediction inflation. When making conditional predictions, people have a certain degree of freedom not only in building a causal scenario leading from the condition to the outcome but also in construing the condition itself. We proposed that merely considering an outcome in conjunction with a condition activates aspects of the condition that are compatible with the outcome, thus inflating the subjective probability of that outcome. This proposition implies that participants typically focus on confirming evidence. Such focus probably affects also the construal of the outcome, as postulated by support theory.

The evidence for the backward-activation account of prediction inflation comes from the results comparing forward-associated and backward-associated pairs. What is special about backward pairs is that the associations from the target to the cue create the feeling that the target is a plausible response to the cue even though its actual probability is very low. In the control condition of Experiment 4, for example, the predicted percentage for the backward pairs averaged about 64%, when the actual percentage was only 2%, that is, an overestimation by a factor of 32. Although prediction inflation was significant for the forward pairs as well, its magnitude was much smaller (a factor of 1.3).

These results are consistent with the backward activation account of prediction inflation. This account has much in common with the selective accessibility account proposed by Strack and Mussweiler (1997; Mussweiler & Strack, 1999) to explain the anchoring effect—the tendency of participants to assimilate a numeric estimate of a target value to that of a previously considered standard. According to Strack and Mussweiler, the presentation of the anchor increases the accessibility of features that the target and anchor have in common, resulting in the assimilation of the final estimate to the anchor value. In a similar manner, the presentation of a condition together with a potential outcome may be assumed to increase the accessibility of features that are common to them.

Examination of the function relating predicted occurrence to actual occurrence (see also Maki, 2005a, 2005b) provides additional insight into the process underlying prediction inflation. In Experiment 1, the strongest inflation occurred for the low-association pairs. Experiments 3A and 3B also yielded results suggesting that participants underestimated the difference between the rate of occurrence of the primary and secondary associates. At the same time, prediction inflation was not inordinately strong for the unrelated pairs. Although there was some overestimation even for the zero-association pairs, this overestimation was slight and may simply reflect the fact that participants could only err in the direction of overestimation. This pattern of results suggests a kind of step function: A certain degree of supportive evidence is sufficient to boost predictions markedly (see also Koriat, 1981).

Nevertheless, it should be stressed that the reported estimates in our study were not completely out of touch with reality: Resolution or discrimination accuracy was quite high (see Figure 3 and Maki, 2005a, 2005b). As Sanbonmatsu et al. (1997) emphasized, the tendency to process evidence in a confirmatory manner may not result in estimation inflation if confirmatory evidence is not available. Thus, in our experiments participants did not avoid reporting a 0 estimate for most of the zero-association pairs. Thus, when supporting evidence is lacking, the tendency for overestimation is low or nonexistent.

Conditional Predictions and Causal Inference

The task used in this study to elicit conditional predictions possibly activates a cause-and-effect construction. Thus, the inordinately marked prediction inflation observed for the backward pairs may also derive from the inversion of conditional probabilities (Ahn & Nosek, 1998). The results of Sherman, McMullen, and Gavanski (1992) suggest that such inversion sometimes occurs when the conditioning information does not correspond to a natural category. When people made estimates that were conditioned on a natural category (e.g., "of 100 randomly selected men, how many prefer blue rather than brown?") the sum of two complementary (i.e., mutually exclusive and exhaustive) estimates was very close to 100%. In contrast, when the sample space did not correspond to a natural category (e.g., estimate gender conditional on preference) this sum departed markedly from 100%, presumably because participants rely on the inverses of the given judgments. It is important to note that this error led not only to overestimation but also to underestimation. Thus, some mechanisms might, perhaps, be expected to result in underestimation.

This observation raises the question whether there are other conditions that may engender underestimation rather than overestimation. In terms of causal-model theories (Waldmann, Holyoak, & Fratianne, 1995) the tasks used in this study (e.g., in Experiments 4, 5, and 6) are likely to be interpreted by participants as involving a common-cause structure in which the same cue produces different effects. It is interesting to inquire whether the overestimation bias is specific to such situations, or whether it is also obtained in assessing the likelihood of a particular cause given a particular effect (Ahn & Bailenson, 1996). For example, it has been observed that causes are perceived to compete with one another in tasks that call for predictive probabilities, unlike effects, which do not seem to compete when diagnostic probabilities are assessed (Ahn & Nosek, 1998; Waldmann & Holyoak, 1992). Thus, perhaps the results of the present study, suggesting lack of complete competition between alternative outcomes (even though in our paradigm these outcomes are mutually exclusive), are specific to predictions from cause to effects and may not generalize to judgments of causes from effects. However, if our account of prediction inflation is correct, we should expect it to hold for judgments from effects to causes as well. In that case, it is the cue that should perhaps bring to the fore aspects of the target in a process similar to that postulated by Tversky and Koehler (1994). Thus, we should expect overestimation whenever the presentation of the condition−outcome pairing activates aspects of the condition and/or outcome that are not transparent when each of them is presented alone. This expectation deserves investigation.

The Overprediction Effect

Experiments 3A and 3B yielded a strong and consistent overprediction that violates the convention that the probabilities assigned to an exhaustive set of mutually exclusive events should add up to one (see Fiedler & Armbruster, 1994; Sanbonmatsu et al., 1997; Teigen, 1974a, 1974b, 1983). In Experiment 3A, the estimated occurrence of the primary and secondary associates together averaged 142.9% despite the fact that each stimulus word elicited many other responses in addition to these associates. Although the effect was somewhat reduced in a within-participant design (Experiment 3B), it was still marked even then.

These results replicate previous findings that were obtained with other paradigms. Some of these were discussed by Tversky and Koehler (1994; see also Dougherty & Hunter, 2003) in terms of the notion that the unpacking of the focal hypothesis increases the judged probability of that hypothesis. Fiedler and Armbruster (1994) also showed inflated estimates when an event is decomposed into several components, but argued that this overestimation derives from the tendency to overestimate low-frequency events. Sanbonmatsu et al. (1997) proposed that overprediction results from confirmatory processes characterizing the selective testing of hypotheses. An interesting finding in their study is that overprediction was found when the evidence favoring each of the judged alternatives was strong. In contrast, when that evidence was weak, an underprediction effect was obtained. As noted earlier, we have some doubt that conditional predictions can also sometimes yield a similar underprediction effect, but this possibility is worth exploring.

The overprediction effect has also much in common with the disjunction effect reported by Tversky and Shafir (1992; see also Bastardi & Shafir, 1998). In one of their experiments, undergraduate students received the following scenario: You have just taken a tough qualifying examination. You have an opportunity to buy a vacation package at an exceptionally low price. The special offer expires before you know your grade. Would you buy the ticket? The majority of students said "no." However, when the scenario was modified to indicate that the student passed or failed the exam, the majority of both fail and pass students said "yes." Tversky and Shafir proposed that this pattern of results occurs when each of the two conditions (e.g., success or failure), activates different reasons for making the same choice. Although in the disjunction effect of Tversky and Shafir the same outcome is assumed to activate different aspects of different conditions that support that outcome, the underlying process may be similar to that which we assume to underlie the overprediction effect, in which different outcomes recruit different supportive aspects of the same condition.

The Effects of Generating One's Own Responses

A most significant finding of this study is our failure to eliminate the overestimation bias. All of our attempts have been basically unsuccessful (see also Maki, 2005a, 2005b). These attempts were inspired by previous findings indicating that overestimation and overconfidence can be effectively reduced by having participants consider or explain alternative outcomes to the one that they are to judge (Koriat et al., 1980; Lord et al., 1984; Hirt & Markman, 1995). The results of Experiments 4, however, indicated that even after generating a response that differed from the subsequently presented target response, participants still overestimated the occurrence of the target response. Although the generation manipulation did reduce prediction inflation for different trials, the effect was quite limited. The generation of two associates to the cue word, either before or after seeing the response (Experiments 5 and 6), was also ineffective in reducing the inflation markedly even when neither of the two associates matched the target response.

Thus, after having had firsthand experience that proved to support alternative outcomes, participants still overestimated the occurrence of the target outcome. This result appears to be at variance with the commonly held view that people estimate probabilities or frequencies by mentally simulating events (Tversky & Kahneman, 1973). In fact, one advantage of the judgment task used in our experiments is that participants can readily simulate the process of producing associations to the cue word. Kelley and Jacoby (1996), for example, obtained results suggesting that participants judge the difficulty of anagrams for others by observing their own experience in attempting to solve these anagrams themselves. Why then did the generation task have such little effect on one's predictions for others when the outcome of the generation process differed from the target outcome?

In our view, this is because in making conditional predictions, people assess the strength of support for one outcome almost independently of support for competing outcomes (Robinson & Hastie, 1985; Sanbonmatsu et al., 1997; Van Wallendael & Hastie, 1990). As discussed in connection with the overprediction effect, the fact that one outcome appears quite plausible does not preclude the possibility that another outcome will also feel very plausible. Because participants focus on confirming evidence, a different response from the participant's own response can reveal aspects of the cue word that are consistent with it, thus largely overriding the experience gained from the generation task.

The question still remains: Why were the effects of the generation task quite limited given the extensive evidence indicating that having participants explain alternative outcomes reduces their confidence in the target outcome (Koehler, 1991)? To answer this question we refer to the distinction between heuristic, experience-based judgments and analytic, information-based judgments (Kelley & Jacoby, 1996; Koriat, Bjork, Sheffer, & Bar, 2004; Koriat & Levy-Sadot, 1999; Strack, 1992). When participants are instructed to explain the occurrence of alternative outcomes or to build a causal scenario for them, they engage in an analytic, inferential process. The results of that process can then help them overcome the inflated experience-based judgments. This analytic process is similar to that which occurs when participants correct their experience-based judgments after realizing that their subjective experience had been contaminated by irrelevant factors (e.g., Gilbert, 2002; Jacoby & Whitehouse, 1989; Strack, 1992). It would seem that the experiential feedback that participants gain from generating their own spontaneous associates can reduce prediction inflation only if participants take advantage of that experience to engage in an analytic, inferential process similar to the process of attempting to provide reasons for alternative outcomes. In that case we might even expect an underprediction bias that results from overcorrection (e.g., Jacoby & Whitehouse, 1989).

Some support for this proposal comes from an exploratory experiment in which we gave 10 participants monetary incentives for making accurate estimates, and at the end of the experiment we asked them how they made their predictions. Three participants stated explicitly that when the cue and target words were related semantically, they found it difficult to ignore the stated target and its association to the cue, although they thought that if they had seen the cue alone they would have probably produced other responses. These reports suggest a conflict between heuristic-driven feelings and analytic-based knowledge (see Denes-Raj & Epstein, 1994; Windschitl & Wells, 1998). Perhaps the small effects that were found for the generation task in the different trials derive from the occasional use of analytic processes to overcome part of the overconfidence associated with experience-based judgments induced by seeing the cue−target pair.

The Mutability of the Representation of the Conditions

In concluding the article we should note that the results presented in this study disclose two characteristics of conditional predictions that, on the face of it, appear to be inconsistent with each other. On the one hand, the overprediction effect discloses the pliability of conditional predictions in that each of two different, mutually exclusive outcomes is perceived to be likely, perhaps with little competition between them. On the other hand, the relative indifference of conditional predictions to the effects of the generation manipulation suggests that conditional predictions are quite resistant to change. In fact, the strongest evidence for the perseverance of conditional predictions comes from the results of the exclusion condition of Experiment 2. These results suggest that once participants are exposed to the target outcome, they find it difficult either to access alternative outcomes or to take them into account in assessing the probability of the target event.

Both of these seemingly incompatible characteristics, however, derive in our view from the two assumptions discussed earlier. First, that in making conditional predictions people focus on confirmatory evidence, with little consideration of alternative outcomes that may result from the same condition. Second, the presentation of a target outcome highlights aspects of the condition that are compatible with it, thus in a sense, modifying the problem itself (see Koehler, 1991).

These two assumptions are necessary to explain the overall pattern observed. On the one hand, they can explain the prediction inflation bias as well as the overprediction effect. On the other hand, they can explain why the effects of the generation manipulations of Experiments 4 and 5 were so limited. They may also explain the ineffectiveness of the exclusion condition of Experiment 2 and of the generation condition used in Experiment 6. In both of these conditions participants were first exposed to the target outcome. In the former case they were asked to estimate the likelihood of an outcome other than the target outcome whereas in the latter case they were asked to produce two other responses before judging the likelihood of the target outcome. We propose that in both conditions the presentation of the cue word together with the target word acts to constrain the judgment task by highlighting those aspects of the problem that support the occurrence of the focal target. Indeed, according to Hoch's (1984) interference hypothesis, whatever people think about first inhibits later retrieval and generation. In his study, participants who first thought of reasons why a particular event might occur in the future found it subsequently difficult to think of reasons why it might not occur, and vice versa. According to Hoch, the difficulty thinking of alternative reasons is used by participants as a cue to infer that alternative outcomes are actually less likely.

Needless to say, more research is needed to specify the processes that contribute to the effects reported in this study. In particular, it is important to specify the extent to which the prediction inflation results from biased evidence search or biased evidence evaluation. In parallel, it is important to determine whether the ineffectiveness of the manipulations designed to reduce prediction inflation (the generation and exclusion conditions) derives from difficulties in accessing alternative outcomes or from difficulties in taking them into account when assessing the probability of the target event. In addition, it is important to specify the contribution of heuristic and analytic processes to the inflation of conditional predictions.

Finally, despite the methodological advantages of the task we used to investigate prediction inflation, this task is clearly remote from many of the real-life contexts in which people normally make conditional predictions. If the results of the present study are to have some implications for real-world decision making, it is critical to show that similar results are also obtained with other ecologically representative predictions.

References

Ahn, W., & Bailenson, J. (1996). Causal attribution as search for underlying mechanisms: An explanation of the conjunction fallacy and the discounting principle. *Cognitive Psychology, 31,* 82–123.

Ahn, W., & Nosek, B. A. (1998, August 15). *Heuristics used in reasoning with multiple causes and effects.* Paper presented at the meeting of the Cognitive Science Society, Madison, WI.

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39,* 1037–1049.

Ayton, P., & McClelland, A. G. R. (1997). How real is overconfidence? *Journal of Behavioral Decision Processes, 10,* 153–285.

Bastardi, A., & Shafir, E. (1998). On the pursuit and misuse of useless information. *Journal of Personality and Social Psychology, 75,* 19–32.

Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology, 14,* 88–96.

Christensen-Szalanski, J. J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes, 48,* 147–168.

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

Dougherty, M. R. P., Gettys, C. F., & Thomas, R. P. (1997). The role of mental simulation in judgments of likelihood. *Organizational Behavior and Human Decision Processes, 70,* 135–148.

Dougherty, M. R. P., & Hunter, J. (2003). Probability judgment and subadditivity: The role of working memory capacity and constraining retrieval. *Memory & Cognition, 31,* 968–982.

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.

Fiedler, K. (2000). On mere considering: The subjective experience of truth. In H. Bless & J. P. Forgas (Eds.), *The message within: The role of subjective experience in social cognition* (pp. 13–36). Philadelphia: Psychology Press.

Fiedler, K., & Armbruster, T. (1994). Two halfs may be more than one whole: Category-split effects on frequency illusions. *Journal of Personality and Social Psychology, 66,* 633–645.

Fiedler, K., Armbruster, T., Nickel, S., Walther, E., & Asbeck, J. (1996). Constructive biases in social judgment: Experiments on the self-verification of question contents. *Journal of Personality and Social Psychology, 71,* 861–873.

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288–299.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90,* 239–260.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 552–564.

Garry, M., & Polaschek, D. L. L. (2000). Imagination and memory. *Psychological Science, 9,* 6–9.

Garskof, B. E., & Forrester, W. (1966). The relationships between judged similarity, judged association, and normative association. *Psychonomic Science, 6,* 503–504.

Gilbert, D. T. (2002). Inferential correction. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 167–184). New York: Cambridge University Press.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin, 107,* 311–327.

Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology, 69,* 1069–1086.

Hirt, E. R., & Sherman, S. J. (1985). The role of prior knowledge in explaining hypothetical events. *Journal of Experimental Social Psychology, 21,* 519–543.

Hoch, S. J. (1984). Availability and interference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 649–662.

Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 719–731.

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 566–581.

Hoffrage, U., & Pohl, R. F. (2003). Research on hindsight bias: A rich past, a productive present, and a challenging future. *Memory, 11,* 357–377.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118,* 126–135.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 49–81). New York: Cambridge University Press.

Kelley, C. M. (1999). Subjective experience as basis of "objective" judgments: Effects of past experience on judgments of difficulty. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 515–536). Cambridge, MA: MIT Press.

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language, 35,* 157–175.

Koehler, D. J. (1991). Explanation, imagination, and confidence in judgment. *Psychological Bulletin, 110,* 499–519.

Koriat, A. (1981). Semantic facilitation in lexical decision as a function of prime-target association. *Memory & Cognition, 9,* 587–598.

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31,* 187–194.

Koriat, A., & Bjork, R. A. (in press). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory & Cognition.*

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General, 133,* 643–656.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 483–502). New York: Guilford.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 107–118.

Levy, A. S., & Pryor, J. B. (1987). Use of the availability heuristic in probability estimates of future events: The effect of imaging outcomes versus imaging reasons. *Organizational Behavior and Human Decision Processes, 40,* 219–234.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (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, England: Cambridge University Press.

Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology, 476,* 1231–1243.

Maki, W. S. (2005a). *Judgments of associative memory.* Manuscript submitted for publication.

Maki, W. S. (2005b, November). Separating bias and sensitivity in judgments of associative memory. Paper presented at the meeting of the Psychonomic Society, Toronto, Ontario, Canada.

Mulford, M., & Dawes, R. M. (1999). Subadditivity in memory for personal events. *Psychological Science, 10,* 47–51.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 35,* 136–164.

Nelson, D. L., Dyrdal, G. M., & Goodmon, L. B. (2005). What is preexisting strength? Predicting free association probabilities, similarity ratings, and cued recall probabilities. *Psychonomic Bulletin & Review, 12,* 711–719.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). *The University of South Florida word association, rhyme and word fragment norms.* Available at http://luna.cas.usf.edu/~nelson/

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2,* 267–270.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175–220.

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin, 125,* 737–759.

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

Rubinsten, O., Anaki, D., Henik, A., Drori, S., & Faran, Y., (2005). Norms for free associations in the Hebrew language. In A. Henik, O. Rubinsten, & D. Anaki (Eds.), *Word norms for the Hebrew language (in Hebrew)* (pp. 17–34). Beer Sheva, Israel: Ben Gurion University of the Negev.

Sanbonmatsu, D. M., Posavac, S. S., & Stasney, R. (1997). The subjective beliefs underlying probability overestimation. *Journal of Experimental Social Psychology, 33,* 276–295.

Sanna, L. J., Schwarz, N., & Stocker, S. L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 497–502.

Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin, 11,* 118–127.

Sherman, S. J., McMullen, M. N., & Gavanski, I. (1992). Natural sample spaces and the inversion of conditional judgments. *Journal of Experimental Social Psychology, 28,* 401–421.

Sherman, S. J., Zehner, K. S., Johnson, J., & Hirt, E. R. (1983). Social explanation: The role of timing, set, and recall on subjective likelihood estimates. *Journal of Personality and Social Psychology, 44,* 1127–1143.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249–275). Hillsdale, NJ: Erlbaum.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73,* 437–446.

Teigen, K. H. (1974a). Overestimation of subjective probabilities. *Scandinavian Journal of Psychology, 15,* 56–62.

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

Teigen, K. H. (1983). Studies in subjective probability III: The unimportance of alternatives. *Scandinavian Journal of Psychology, 24,* 97–195.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: A handbook of basic principles* (pp. 239–270). New York: Guilford Press.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 359–380.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 4,* 207–232.

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

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 3,* 305–309.

Van Wallendael, L. R., & Hastie, R. (1990). Tracing the footsteps of Sherlock Holmes: Cognitive representations of hypothesis testing. *Memory & Cognition, 18,* 240–250.

Wänke, M., Bless, H., & Biller, B. (1996). Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin, 22,* 1105–1113.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121,* 222–236.

Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (1995). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General, 124,* 181–206.

Wilson, T. D., & LaFleur, S. J. (1995). Knowing what you'll do: Effects of analyzing reasons on self-predictions. *Journal of Personality and Social Psychology, 68,* 21–35.

Windschitl, P. D., & Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology, 75,* 1411–1423.

Winman, A., & Juslin, P. (2006). "I'm *m/n* confident that I'm correct": Confidence in foresight and hindsight as a sampling probability. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 409–439). Cambridge, England: Cambridge University Press.

Yaniv, I., Yates, J. F., & Smith, J. E. K. (1991). Measures of discrimination skill in probabilistic judgment. *Psychological Bulletin, 110,* 611–617.

Received October 27, 2005
Revision received March 13, 2006
Accepted March 15, 2006 ■