Cognitive Illusions, Heuristics, and Climate Prediction

Neville Nicholls
Bureau of Meteorology Research Centre, Melbourne, Australia

ABSTRACT

A number of studies in meteorological journals have documented some of the constraints to the effective use of climate forecasts. One major constraint, the considerable difficulty people have in estimating and dealing with probabilities, risk, and uncertainty, has received relatively little attention in the climate field. Some of this difficulty arises from problems known as cognitive illusions or biases. These illusions, and ways to avoid them impacting on decision making, have been studied in the fields of law, medicine, and business. The relevance of some of these illusions to climate prediction is discussed here. The optimal use of climate predictions requires providers of forecasts to understand these difficulties and to make adjustments for them in the way forecasts are prepared and disseminated.

1. Introduction

More than two decades ago, Glantz (1977), in a study of the value of drought forecasts for the African Sahel, noted that a variety of social, economic, environmental, and political constraints would limit the value of even a perfect drought forecast, at that time. A few years later Lamb (1981) questioned whether we even knew what sort of climate variables we should be predicting. He pointed out that the achievement of the ultimate goal of improved climate prediction—the reduction of the adverse socioeconomic consequences of climate variability—had two prerequisites. First, the activities and regions most impacted by climatic variations required identification. Second, the regional economies with the flexibility to adjust to capitalize on the availability of skillful climate forecasts needed to be determined. Only after these prerequisites had been met could we focus on the development of climate prediction systems likely to result in real benefits.

Corresponding author address: Dr. Neville Nicholls, Bureau of Meteorology Research Centre, P.O. Box 1289K, Melbourne, VIC 3001, Australia.
E-mail: n.nicholls@bom.gov.au
In final form 5 April 1999.
©1999 American Meteorological Society

Since these two papers were published there have been substantial advances in our understanding of parts of the climate system, especially the El Niño–Southern Oscillation, which provides some climate predictability in many parts of the world (Trenberth 1997). There has also been research aimed at determining the value of climate forecasts, in a variety of settings (e.g., Hulme et al. 1992; Adams et al. 1995; Hammer et al. 1996; Mjelde et al. 1997; Pulwarty and Redmond 1997).

There have also been a few studies aimed at identifying why potential users were *not* incorporating climate predictions in the decision-making processes. Changnon et al. (1995) surveyed decision makers in power utilities, to discern the needs and uses of climate forecasts. Only 3 of the 56 decision makers surveyed used forecasts. Major hindrances to the use of forecasts included the following: forecasts difficult to interpret, lack of models to integrate information, uncertainty over accuracy, additional information necessary, proof of value necessary, lack of access to expertise, difficult to assess forecasts.

A similar list of difficulties was assembled by Pulwarty and Redmond (1997), to account for the lack of use of climate forecasts involved in salmon management issues in the Columbia River basin, despite the significant influence of precipitation and its subsequent hydrologic impacts on the region. The reasons

advanced by decision makers here included the following: forecasts not "accurate" enough, fluctuation of successive forecasts, what is the nature of the forecast, external constraints (e.g., legal requirements) forbid a flexible response to forecast information, procedures for acquiring knowledge and implementing decisions by incorporating forecasts have not been defined, availability of locally specific information may be more important, "value" may not have been demonstrated (by a credible organization or competitor), required information may not have been provided, competing or conflicting forecast information, lack of information regarding how well the forecasts are "tracking" the actual climate, history of previous forecasts not available.

Hulme et al. (1992), in a study of the potential use of climate forecasts in Africa, concluded that rural communities (e.g., individual farmers) were unlikely to obtain direct benefits from forecasts. This was because of the following.

- The forecast information is not precise enough to influence local decisions such as where to plant or where to move animals.
- All producers are restricted in their flexibility to respond to forecast information; the poorer and more vulnerable the producer, the greater the restriction.
- Decisions are based on a range of factors of which climate is only one. A climate forecast in isolation from other information is unlikely to improve on existing knowledge systems.

Most of these impediments to the use of climate predictions represent a lack of knowledge of either the forecast system or the impact system, or of forecast delivery problems, or of difficulties in users reacting to the forecasts. Nicholls (1999) discusses the sources of many of these impediments and the problems they cause. Another group of impediments to the correct use of forecasts exist: cognitive illusions. These are analogous to optical illusions in leading to errors we commit without knowing we are doing so, except they arise from our difficulties in quantifying and dealing with probabilities, uncertainty, and risk. They affect people of all levels of expertise and field (with the possible exception, in some cases, of statisticians). They lead to departures from "rational thinking," defined by Baron (1994) as "whatever kind of thinking best helps people achieve their goals."

Slovic (1987) summarizes the nature and effects of these illusions.

> Research on basic perceptions and cognitions has shown that difficulties in understanding probabilistic processes, biased media coverage, misleading personal experiences, and the anxieties generated by life's gambles causes uncertainty to be denied, risks to be misjudged (sometimes overestimated and sometimes underestimated), and judgments of fact to be held with unwarranted confidence. Experts' judgments appear to be prone to many of the same biases as those of the general public, particularly when experts are forced to go beyond the limits of available data and rely on intuition. . . . Strong initial views are resistant to changes because they influence the way that subsequent information is interpreted. New evidence appears reliable and informative if it is consistent with one's initial beliefs; contrary evidence tends to be dismissed as unreliable, erroneous, or unrepresentative. . . . When people lack strong prior opinions, the opposite situation exists—they are at the mercy of the problem formulation. Presenting the same information about risk in different ways (for example, mortality rates as opposed to survival rates) alters people's perspectives and actions.

The difficulties people have in dealing with probabilities and uncertainties, as summarized by Slovic, have clear implications to attempts to have climate predictions (which are inherently uncertain and probabilistic) used in an optimal fashion. Little recognition appears to have been given to the problems cognitive illusions may cause users (and providers) of climate predictions.

These cognitive illusions do not simply reflect general ignorance of probabilities. There is a common belief among meteorologists that the public is largely ignorant about probabilities in general and their appropriate use. Konold (1989) found that interview subjects had considerable difficulties correctly interpreting statements such as "70% chance of rain." There has been some work on the difficulties people have in interpreting probabilities in weather forecasts (Murphy et al. 1980; Sink 1995) and how to write forecasts in such a way as to reduce user confusion (Vislocky et al. 1995). Some of the difficulty arises from what Murphy (1996) calls "event misinterpretation" (e.g., misinterpreting the event "precipitation," rather than misinterpreting the probabilities). Baker (1995) found that "people are more capable of comprehending and using at least certain types of probability information than is often assumed." The cognitive problems noted

by Slovic, and discussed below, are more subtle than just a general ignorance of probabilities. They are more fundamental to the way we (i.e., forecast providers as well as users) deal with uncertain situations and require careful thought and research about how we should write, deliver, and value climate predictions.

2. Cognitive illusions

Some cognitive illusions arise because the capacity of the human mind for solving complex problems is limited, compared with the size of problems whose solution is required for objectively rational behavior in the real world. So, people use simple rules of thumb or heuristics to simplify decision making. In general, heuristics are helpful, but they do lead to biases in many situations, especially in uncertain situations where probabilities are encountered. Some other problems in dealing with uncertainty reflect our inability to recognize factors that appear to be counterintuitive—some of these are easily overcome, once the problem is identified.

The following are some of the well-known sources of cognitive bias. Many of these situations and problems have been studied extensively in the economic, legal, and medical area. Much of the work on these problems stems from work by psychologists A. Tversky and D. Kahneman in the early 1970s (Tversky and Kahneman 1974, 1981; Kahneman and Tversky 1982), although Armstrong (1999) cites much earlier research pointing to some of these issues. Piattelli-Palmarining (1994), Plous (1993), and Bazerman (1994) provide introductions to cognitive illusions and decision making, and many of the examples here are drawn from them. Most of the examples described here, however, are from laboratory experiments. Stewart (1997) has urged caution in generalizing from these studies to decision making in a real-world context. I have tried, where possible, to illustrate how these problems may have affected the use and preparation of climate forecasts during the 1997/98 El Niño, to demonstrate their potential impact in real-world decision making with climate forecasts.

There is considerable overlap between the different sources or types of bias or illusion, and it is often difficult to determine, for a particular example, which mechanism is responsible for the observed bias. So, some of the climate examples given in the following, and suggested as arising from a specific bias, could also reflect one or more other sources of bias.

a. Framing effect

The way a problem (or forecast) is posed can affect a decision. As an example, imagine that New York faces an outbreak of an unusual disease that is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Which program would you favor? If A is adopted, 200 people will be saved. If B is adopted there is one-third probability that 600 people will be saved and a two-thirds probability that nobody will be saved. Tests indicate that a majority of respondents (72%) would select program A, the risk-averse option.

What if the wording is changed slightly? Now, which of these two programs would you favor? If C is adopted, 400 people will die. If D is adopted, there is a one-third probability that nobody will die, and two-thirds probability that 600 people will die.

In this case, a clear majority (78%) of respondents prefer program D, the risk-taking option. Even though programs A and C are essentially identical, as are programs B and D, a slight change in the "framing" of the questions leads to a substantial shift in decision making. Such reversals have been observed in studies of many groups, including university faculty and physicians. Nor is this framing effect simply verbal "trickery." Even when the inconsistency of their selections were pointed out, people often stand by their contradictory choices (Dawes 1988, 36).

The above question may seem rather abstract. There is, however, considerable evidence that framing is a very severe impediment to rational decision making and that even professionals experienced in decision making are still affected by framing. For instance, if doctors were told there is a mortality rate of 7% within five years of a certain operation, they hesitated to recommended it to their patients. If they were told it had a survival rate after five years of 93%, they were more inclined to recommend it to their patients.

In climate prediction, this framing effect suggests that forecasts expressed in terms of the likelihood of drought may lead to different decisions to forecasts expressed as the *nonlikelihood* of wet conditions. For instance, predicting a 30% chance of drought is likely to cause a different response to a forecast of a 70% chance of normal or wet conditions, even though the objective content of the forecast is the same.

Even a change from a worded forecast to a numerical forecast can affect the response by users. For instance, Sink (1995) found that 11% of respondents would rate a forecast "rain is likely" as poor if it did not rain on their house. Of the same respondents, 37%

would rank a forecast "70% chance of rain" as poor if it did not rain on their house, even though they associated the word "likely" in a forecast with a probability of 70%. This suggests that, despite respondents equating 70% and "likely," they still reacted to the worded forecast quite differently to the numerical prediction.

The framing problem can be amplified because people react differently, asymmetrically, to potential gains or losses. For instance, assume you are given a bonus of $300. You may then choose between receiving $100 for sure or tossing a coin. If you win the toss you win an additional $200; if you lose you receive nothing. There is a strong preference for the first (sure) gain.

In the next example you are first given a bonus of $500. You may then choose between losing $100 for sure or tossing a coin. If you lose you must return $200; if you win you do not pay anything. This time, the majority preference is for the gamble.

There is a strong asymmetry between decision making in the face of losses and in the face of gains; we tend to be conservative when offered gains and adventurers when we face loss. The threat of a loss has a greater impact on a decision than the possibility of an equivalent gain (Kahneman and Tversky 1982). Such an asymmetry will clearly affect decision making by those affected by climate variations. Yet studies of the value of climate forecasts do not take such asymmetry into account. This example again illustrates the importance of "framing" the question or forecast. Very different responses can be initiated by small changes in wording of a forecast.

b. Availability

Which of the following causes more deaths in the United States each year: stomach cancer or motor vehicle accidents? Most respondents select motor vehicle accidents, but stomach cancer causes twice as many deaths. The "availability" of media stories about motor vehicle deaths biases our perception of the frequency of events. Similarly, the high profile of the 1982/83 and other intense El Niño events in media stories, and often mentioned in forecasts, may bias a farmer's (or other user's) expectations of an El Niño to the 1982/83 experience. This appeared to happen in Australia during the 1997/98 event: many stories about the El Niño and the forecast situation for 1997 mentioned the disastrous effects of the 1982/83 El Niño, without indicating that other El Niño events have had less deleterious consequences. We are particularly prone to overestimating unlikely events with vivid impacts. Slovic and Fischhoff (1977) discuss the effect of this "availability" heuristic on perceived risks of nuclear power. They point out that any discussion of the potential hazards, regardless of likelihood, will increase the memorability of these hazards, and increase their perceived risks. A similar situation probably affects public perception of the risks of serious impacts from an enhanced greenhouse effect, or of an El Niño forecast.

In Australia, the prevalence of media stories relating El Niño to serious drought (e.g., the 1982/83 situation) now produces a public expectation that any predicted El Niño will lead to severe drought everywhere in eastern Australia. The public are likely to be surprised if an El Niño does not lead to a severe drought, or if a drought occurs without an El Niño. Yet both of these have happened in the past. The availability of information relating to the coincidence of El Niño and drought biases public perceptions of how closely these phenomena are linked. During 1997/98 there were parts of Australia in drought, but because *severe* drought was not widespread, people were surprised.

c. Anchoring and adjustment

Most people, when asked to estimate a numerical value, will start with a number that easily comes to mind and then adjust from that number. Numbers totally irrelevant to the number to be estimated can act as anchors. People usually then struggle to adjust sufficiently away from the initial anchor. For example, Russo and Schoemaker (1989) asked subjects, "What are the last three digits of your home phone number?" They then added 400 to this number and asked "Do you think Attila the Hun was defeated in Europe before or after that year?" Then, without telling the subjects if their answer was correct, they asked, "In what year would you guess Attila the Hun was defeated?" The correct answer to this is A.D. 451. The estimates, however, varied depending on the initial anchor (i.e., the subjects' phone number!), as in Table 1.

Anchoring could also affect recipients of climate forecasts. For instance, mention of the impact of the 1982/83 El Niño (perhaps the percentage reduction i n the Australian wheat yield caused by that event, or the global cost of that event) could anchor a user who is aware of the forecast of an El Niño in 1997/98 to the severity of the 1982/83 event. Then, even if other evidence or information is provided indicating that the 1982/83 event was especially severe and that

the 1997/98 El Niño would be unlikely to result in such a severe and widespread drought, a user of the 1997/98 forecast may still not have adjusted sufficiently from the 1982/83 "anchor" of a very severe drought.

Pielke (1999) notes that Grand Forks residents in the 1997 Red River flood "anchored" to a specific, numerical, categorical forecast of flood height. Pielke argues that the use of probabilities, rather than categorical forecasts, would have reduced this "anchoring."

d. Underweighting base rates

The neglect of prior probabilities in judging the probability of events, the base-rate error, is a common bias. Goodie and Fantino (1996) discuss the following classic example to illustrate the dangers of ignoring base rates.

- A taxi was involved in a hit and run accident at night. Two taxi companies, the green and the blue, operate in the city.
- Eighty-five percent of the taxis are green and 15% are blue.
- A witness identified the taxi as blue.
- The witness identifies the correct color 80% of the time and fails 20% of the time.
- What is the probability that the taxi was blue? This can be determined through Bayes's theorem. It is more common for the witness to see a green taxi and mistakenly call it blue ($0.85 \times 0.20 = 17\%$ of all cases) than for the witness to see a blue taxi and label it correctly ($0.15 \times 0.80 = 12\%$ of all cases). If the witness reports seeing a blue taxi, the probability that the cab actually was blue is $0.12/(0.17 + 0.12) = 41\%$. Subjects responding to this problem, however, ignore the base rates (how many blue and green taxis operate in the city) and rely on the reliability of the witness.

Should your decision on going for a walk or carrying an umbrella rely on the daily weather forecast (Matthews 1996)? This can also be determined with Bayes. The accuracy of the United Kingdom 24-hour rain forecast is 83%. The climatological probability of rain on the hourly timescale appropriate for walks is 0.08 (this is the base rate). Given these values, the probability of rain, given a forecast of rain, is 0.30. The probability of *no* rain, given a forecast of rain, is 0.70. So, it is more likely that you would enjoy your walk without getting wet, even if the forecast was for rain tomorrow.

Clearly, similar problems will appear in climate forecasting as in daily weather prediction. Imagine that a climate prediction model (whose accuracy we estimate at 90%) predicts that my farm will be in drought next year. Imagine also that historically there is a 10% chance of being in drought. Assume further that the model is unbiased, that is, over the long run it forecasts just as many droughts as occur in reality (10%). We can then show that the chances that next year will be a drought year, given that the model forecasts a drought, is only 50% (again from Bayes's theorem). Despite the very high accuracy of the model, the probability of a drought still comes down to a coin toss. Most users, given a forecast of drought and a quoted accuracy of 90%, would most likely use this number in any (subjective) decision making process. It would be difficult to convince users that (a) your model was correct 90% of the time, but that (b) the probability that there will be a drought next year was only 50%, given that the model forecast a drought.

e. Overconfidence

Which of these causes of death is most frequent (and how confident are you?) All accidents, or heart attacks? Homicide, or suicide? Breast tumor, or diabetes? The first alternative in each pair is selected with great confidence by most respondents. The correct answer is the second in each pair.

People tend to be overconfident in their estimates (e.g., of the world population in 1998), or of their answers to questions of this sort. Fischhoff et al. (1977) found that people who assigned odds of 1000:1 of being correct in their estimates were correct less than 90% of the time. For odds of 1 000 000:1 their answers were correct less than 96% of the time. Cerf and

TABLE 1. Estimates of date of Attila's defeat depend on initial anchor (after Russo and Schoemaker 1989).

Range of initial anchor (last three digits of phone number plus 400)	Average estimate of year of Attila's defeat
400–599	629
600–799	680
800–999	789
1000–1199	885
1200–1399	988

Navasky (1998) provide many overly confident (and wrong) quotes from experts, on subjects as diverse as share market crashes, heavyweight boxing match results, and heavier-than-air flight. Even in an area of specialist knowledge, experts regularly overstate confidence. Russo and Schoemaker (1989) quote the results of several studies reporting the overconfidence of experts. For example, employees of a chemical company were asked questions regarding the chemical industry and asked to specify 90% confidence ranges on their answers. The correct answers lay outside these ranges about 50% of the time (instead of 10% of the time). There is evidence that one reason for inappropriately high confidence is our tendency not to search for reasons why we (or a forecast) might be wrong (Baron 1994). Subjects forced to think of reasons why they could be wrong tended to be less overconfident.

Sniezek and Henry (1989) found that groups are just as susceptible to unreasonably high levels of confidence in their judgements. Some 98% of subjects believed that their group judgments were in the top half of all group judgments with respect to accuracy. In the climate prediction situation this suggests that a group working to derive a consensus forecast from several forecast systems may well overestimate the confidence in the group forecast. This would apply in cases when a consensus seasonal forecast is being prepared, but also when "experts" are asked to estimate the probability of a serious impact from an enhanced greenhouse effect.

In early 1998 some climate forecasters were indicating that a La Niña event during 1998 was "almost certain." This seems an example of overconfidence. Arkes (1999) notes that overconfidence is a common finding in the forecasting research literature and suggested that listing reasons why the forecast might be wrong, considering alternative forecasts, and including a devil's advocate in group discussion would help avoid overconfidence. During 1997 many Australian farmers seemed overconfident, in the face of the El Niño and continuing dry conditions, with some asserting that they were not worried because they "always" get good rain by midwinter.

A related problem is optimism about personal risk. People usually claim that they are less likely to be affected by a hazard than are their peers (Weinstein 1989). For instance, a random sample of adults, asked to compare their risks from several hazards with their peers, yielded the following ratios of "below average" to "above average" responses:

- asthma 9:1
- drug addiction 8:1
- food poisoning 7:1
- influenza 3:1
- lung cancer 2:1
- pneumonia 5:1

An optimistic bias was found for 25 of 32 hazards in this study.

If such overoptimism translated to the climate prediction field, some users may judge that their farm (or their decision making) is less vulnerable to the effects of drought than is the case with their neighbors. Such optimism can have positive effects, but it can also skew decision making. In Australia in mid-1997, some farmers were denying that the El Niño affected their farm, but would then note that the worst year they had experienced was 1982, a severe El Niño. This may have reflected lack of knowledge. However, at the time (1997) there was a lot of publicity about the 1982 El Niño, so it may also reflect, to some degree at least, overoptimism. Such farmers were, presumably, less willing to take action based on the forecasts of an impending El Niño. Overoptimism could also lead a user to assume that they were well prepared for El Niño impacts, or at least better prepared than their neighbors. During this period, some Australian farmers operated, apparently, in the hope that despite forecasts of an increased likelihood of drier than normal conditions, their farm would receive good rains. In these cases, documented through media reports, overoptimism is clear.

f. Added information bias

We usually assume that increasing the amount of information (e.g., the number of forecasts from different models) should help potential users. However, it is difficult for users to assimilate such multiple sources of information correctly. Russo and Schoemaker (1989) reported a study of the effect of increasing information available to race track handicappers on the accuracy and confidence of the handicappers. The handicappers' task was to judge the likelihood of each horse winning. To do this they were provided with information regarding the past performance of the horses. Increasing the amount of information from a small amount to a very large amount did not increase accuracy of their handicapping, even though it increased their confidence that their forecasts were correct. Redelmeier and Shafir (1995) reported on a study designed to reveal how doctors respond as the num-

ber of possible options for treatment is increased. The introduction of additional options (cf. different forecasts) raised the probability that the doctors would choose to either adopt the first option or decide to do nothing. So the additional information (in the climate case, extra forecasts) can distort decisions. DeBondt and Thaler (1986) found that when new information arrives, investors revise their beliefs (in this case about stock market performance) by overweighting the new information and underweighting the earlier information. A similar problem (overweighting of newly arrived information such as new forecast) would likely affect the use of climate forecasts. Stewart et al. (1992) argued that weather forecasts improved little when increased information was provided because the information increased the complexity of the forecast task. In the El Niño situation, increasing information regarding the likelihood of drought, from various forecasters, may not be well assimilated by forecast users. Since information and forecasts are now available from many sources, it is important that users assimilate this extra information appropriately.

g. Inconsistent intuition

Many decisions are reached by intuition, after a data gathering exercise. Many farmers, for instance, would decide intuitively how to change their farm management in the face of a predicted drought. Russo and Schoemaker (1989) described an experiment to test how well intuition worked. They asked students to predict the grades of applicants to an MBA course, based on data on each applicant such as work experience, college entrance test scores, etc. They then conducted regression analyses of each student's predictions, using the applicants' attributes as predictors. This regression analysis was then used to develop a new set of predictions. So, the intuitions of the students were used to develop a consistent model. These "bootstrapped" predictions were better at predicting the applicants' scores than were the students, for 81% of the students. Similar results have been documented for predictions such as medical diagnosis, changes in share prices, business failures, and others. Substituting a simple but consistent model for a human's intuition consistently results in improved prediction. This is because the human subjects, even if they have an intuitive model (which could be revealed by the regression analysis), did not apply it consistently.

There is extensive evidence that decisions made with quantitative models routinely outperform decisions made intuitively. Dawes et al. (1989) and Grove and Meehl (1996) provide examples in clinical medicine and psychiatry, and note that metaanalysis of studies comparing intuitive (clinical) and quantitative decision making in social sciences and medicine found that the clinical methods outperformed the quantitative methods in only a small percentage of studies. Despite the weight of this evidence, it appears not to have greatly changed decision making. Dawes et al. (1989) suggest that this may be due partly because the availability of a clinician's own experiences (he or she may recall dramatic instances in which his or her interpretations proved correct) fosters the misappraisal. The other biases discussed in this paper may also contribute to the failure of intuitive decision making.

The evidence from these other fields suggests that intuitive decision making of how to use climate predictions will be inferior in skill relative to objective techniques. As well, intuitive efforts to combine various forecasts, by either users or groups of forecasters, seem likely to produce suboptimal combinations. However, even objective forecast systems can be unreliable. Stewart (1999) notes that, even though statistical models often outperform human judges/forecasters, and that models of the judges often outperform the judges themselves, the analytical models can produce catastrophic errors at times.

h. Hindsight and confirmation bias

Jevons (1958) observed that "men mark where they hit, and not where they miss." After finding out whether or not an event occurred, individuals tend to overestimate the degree to which they would have predicted the correct outcome. Reported outcomes seem less surprising in hindsight than in foresight. Fischhoff (1975) asked five groups of subjects to read a passage about the 1814 British–Gurka war. One group was *not* told the result of the war. The other four groups were told either that the British won, the Gurkas won, a stalemate was reached with no peace settlement, or a stalemate was reached with a peace settlement. Each subject was then asked what his or her subjective assessment of each of the outcomes would have been without the benefit of knowing the reported outcome. The strong finding in this and similar studies is that knowledge of an outcome increases an individual's belief about the degree to which he or she would have predicted that outcome without the benefit of that knowledge.

Two groups of subjects were told about the following fictitious experiment (Slovic and Fischhoff 1977). A team of government meteorologists had recently

seeded a hurricane. Subjects estimated the probability that the result of the seeding (increased intensity or decreased intensity) would be replicated in all, some, or none of six additional hurricanes yet to be seeded. One group was told that the initial hurricane had increased in intensity—the group's average probability that the remaining six would also increase was 0.56. The other group was not told the result of the first seeding, but was asked to estimate the probability that all six additional hurricanes would intensify if the initial hurricane had intensified—the group average probability was 0.31 in this case. There is no reason to expect that the first group (with "hindsight") should be more confident than the second group, other than the existence of this hindsight bias.

Hindsight bias leads to increased confidence in judgment after the fact and can lead to use of faulty decision strategies. The general view, among scientists, of the accuracy of forecasts of the El Niño of 1997 appears to illustrate hindsight bias. None of the climate forecast systems reported in the December 1996 National Oceanic and Atmospheric Administration Experimental Long-Lead Forecast Bulletin predicted anything more than slight warming. Only one model was reported in the March 1997 bulletin as predicting more than slight warming by June–August 1997, and this model's prediction was for a rather moderate El Niño, rather than very strong. Yet the forecasting of the 1997 El Niño with large models is now regarded as "a stunning success" (Kerr 1998). Certainly such models appear to have done quite well in prediction, once the El Niño was established, as did some simple statistical methods. However, the general view of the "outstanding success" seems likely to ensure that model forecasts will be assumed to be of high accuracy in the future.

Some of the model predictions for the 1997/98 event were very poor. It seems these misforecasts are being ignored when the stunning success of the El Niño model forecasts is assessed. There is considerable evidence that people tend to ignore, and not even search for, disconfirming information of any hypothesis (Bazerman 1994), a process called confirmation bias. People seek confirmatory evidence and avoid the search for disconfirming evidence. Wason (1960), in one of the first demonstrations of this tendency, told subjects that he had a rule that classified sets of three integers, and that the triple "2 4 6" conformed to this rule. The subjects were then asked to determine the rule by generating triples of their own. For each triple the experimenter told the subject whether it conformed to the rule. Subjects found the task surprisingly difficult. In general, subjects tested their hypotheses by testing only positive examples (i.e., examples that confirmed with their hypothesis), with the consequence that they failed to receive disconfirming evidence. As a result, most subjects announced at least one incorrect hypothesis.

This confirmation bias has also been detected in researchers themselves. Rosenthal and Fode (1963) led 12 experimenters to believe they would be working with "maze bright" or "maze dull" rats in a maze-running task. In reality the labels were assigned at random. Each experimenter was given five rats and asked to record how they performed on 50 trials. The maze-bright rats substantially outperformed the maze-dull rats.

Confirmation bias has clear relevance to climate prediction. If a user (or forecaster) believes a specific model provides good forecasts, evidence of model inaccuracy may be ignored or not even sought. This factor may also impinge on the climate modeling community; the belief that large coupled models now outperform simple models may lead researchers to play down cases where the simple models or statistical methods outperformed the large models. In the assessment of the 1997/98 forecasts, it appears that only the better (in hindsight) forecasts are being considered—the poor forecasts from some models are largely ignored.

i. Belief persistence: Primacy and inertia effects

Primacy and inertia also tend to weight evidence inaccurately. Many psychological studies have reported primacy effects in judgment (see Baron 1994, 283–286)—people tend to weight more heavily evidence presented first. Asch (1946) gave subjects a list of adjectives describing a person, such as "intelligent, industrious, impulsive, critical, stubborn, envious" or "envious, stubborn, critical, impulsive, industrious, intelligent." The impressions of the person were more favorable given the first list than the second, even though one list is just the reverse of the other. It seems this sort of effect might bias a user's perception (or a forecaster's) perception of a forecast system. If the first forecast produced or used is correct, the user (or forecaster) may overestimate the reliability of the system. Forecast producers, with forecasts from several models, may weight the first forecast received more heavily than those received later.

Inertia may lead people to ignore evidence that contradicts their prior belief (e.g., that a particular forecast system produces useful forecasts). Lord et al. (1979) gave subjects (who had previously indicated

that they favored, or opposed, capital punishment) two reports, one purporting to show that capital punishment was effective in deterring crime, the other purporting to show that it was ineffective. This balanced evidence resulted in subjects' beliefs becoming stronger, regardless of whether they favored or opposed capital punishment. If anything, the mixed evidence should have made the subjects less sure of their belief. In the climate prediction area, this suggests that users who believed in the utility of forecasts would have their beliefs reinforced even by mixed evidence, thereby overweighting cases where the forecasts were accurate, and underweighting forecast failures. Forecast producers may not include in their forecast preparation sufficient recognition of the disparity of model predictions, relying instead too heavily on a forecast that supported their intuitive understanding of the current state of the climate.

j. Group conformity and decision regret

The pervasiveness of "groupthink" is well documented (Janis 1982). Many decisions taken in groups are erroneous. Cohesiveness, insulation, and stress generally lead groups to reach consensus too quickly, often supporting whatever was initially advocated by a group leader. Groups then focus almost entirely on information confirming the decision, rather than searching for disconfirming evidence. Even small "groups" without obvious leaders are prone to fall into this trap. For instance, Russo and Schoemaker (1989) discuss the example of the "Asch" test. Each subject was asked whether the test line in Fig. 1 was equal in length to line A, B, or C.

- If subjects were tested individually, 99% of subjects answered B.
- If person in front of the test subject said A, the error rate increased from 1% to 2%.
- If two people ahead of the test subject said A, the error rate increased to 13%.
- If three people ahead of the test subject said A, the error rate increased to 33%.
- If, as well, the subject was told a monetary reward for the group as a whole depended on how many members of the group gave the correct answer, the error rate increased to 47%.

Gigone and Hastie (1997) found that small groups undertaking judgment (forecast) tasks were less accurate than the mean judgment of the individuals making up the groups. The groups were more accurate than most of the individuals, but this was simply due to averaging (and the reduction in random errors this averaging produces) and the group dynamics did not add anything more than this averaging process. The groups tended to make more variable (extreme) judgments than the means of the individual judgments and this reduced the group accuracy. This increased variability may reflect a tendency for groupthink.

Seasonal climate predictions are often prepared in small group discussions. In such discussions, the possibility of groupthink biasing the forecasts would need to be considered. As well, many users of forecasts (e.g., farmers) operate in small groups and discuss climate predictions within these groups. Again, groupthink may lead to biased decision making in this situation.

A related problem, decision regret, can arise from perceived pressure to continue to use a conservative strategy, to avoid ridicule from peers. Keynes, in a discussion about an unconventional investor, noted that "if his decisions are unsuccessful . . . he will not receive much mercy. Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally" (Keynes 1935). This is a form of decision regret. In the case of climate forecasts, imagine a farmer who has a drought forecast available to him and consider these two combinations of action and outcome.

- He chooses to ignore the forecast and continue his normal farm management techniques and the drought occurs.

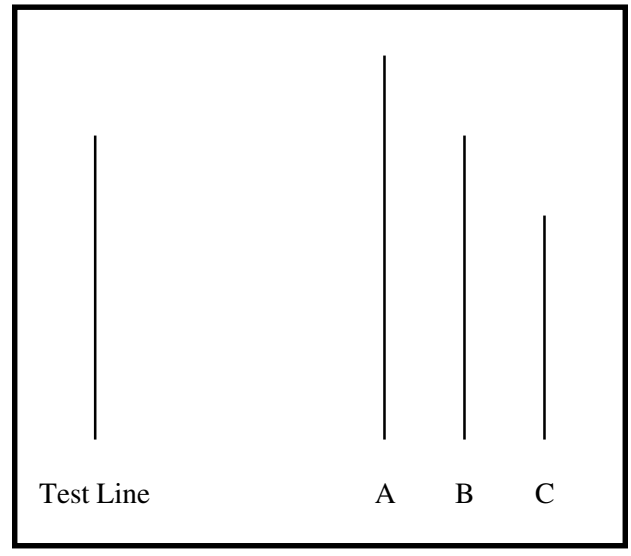

FIG. 1. The "Asch" test. Is the test line equal in length to A, B, or C?

- He alters his farming strategies based on the forecast and the drought does not occur.

Which combination would cause the farmer more regret? Decision regret theory (cf. Keynes) would suggest that adopting the unconventional action (changing the farming practice because of the drought forecast) would cause more regret. The regret associated with a loss incurred by an action tends to be more intense than the regret associated with inaction or a missed opportunity (Kahneman and Tversky 1982). This appears to have occurred in Australia during the 1997 El Niño. Some farmers took action in anticipation of a severe drought. The drought did not eventuate and many farmers expressed their anger through the media (Nicholls and Kestin 1998). This factor (decision regret) would tend, one imagines, to cause farmers to not react to climate predictions. Of course, if it became the usual practice for farmers to use climate predictions, then ignoring the forecast will become the unconventional (and therefore unlikely) practice.

3. Discussion

I have here highlighted just some of the many problems individuals and groups encounter in dealing with uncertainty. Plous (1993) discusses even more sources of bias and apparently irrational decision making. The existence of so many well-documented sources of bias means that individuals seem unlikely to realize the full benefits of seasonal climate prediction, without considerable efforts to overcome these biases. Even those of us involved in preparing and disseminating forecasts need to be aware of the likelihood that our cognitive processes could lead to biased predictions.

Much of the above dealt with problems people have in dealing with probabilities. This discussion may lead some to the conclusion that issuing forecasts in categorical terms, rather than probabilities, would be less likely to cause confusion and misuse. There are, however, strong arguments for the inclusion of probabilities in forecasts, to indicate the expected accuracy (e.g., Vislocky et al. 1995; Fischhoff 1994) and to reduce "anchoring" (Pielke 1999). As well, as noted earlier, there is evidence that it is not the probabilities that cause the confusion (Baker 1995; Murphy 1996). Hence the task is to ensure that the forecasts and probabilities are written in such a way that these cognitive biases do not lead to incorrect decision making on the part of forecast users. The real issue is in the structuring of the communication and decision processes to prepare and disseminate the information rather than the information per se.

Bazerman (1994) outlines four strategies to improve decision making and avoid these biases and paradoxes.

1) Decision makers need to develop a conceptual understanding of rational decision making and to recognize the biases that limit the rationality of decision making.
2) Debias judgment, through "unfreezing" and changing decision-making processes.
3) Use linear models based on expert judgment to make decisions (e.g., the bootstrapping example discussed above).
4) Adjust intuitive decisions to account for bias (for instance, by recognizing the framing effect that a particular forecast might produce).

All of these strategies are difficult to apply, and the evidence for their success is rather weak. However, without such strategies it seems likely that climate forecast providers and forecast users will continue to indulge in incorrect, biased decision making. Without at least a recognition that decision making is inherently biased, we are unlikely to realize the benefits of improved climate forecasts. At the very least, greater efforts are needed to develop quantitative decision-making schemes for the use of climate predictions in a variety of settings. This is much more than the development of demonstration projects of how to apply forecasts; schemes to allow the general application of the forecasts, across a wide variety of contexts and decisions, are needed. As well, greater care is needed to understand how written forecasts, and numerical probability forecasts, are interpreted and understood by potential users. There has been some work done in this area for daily weather forecasting (e.g., Murphy and Brown 1983a,b; Fischhoff 1994; Sink 1995), but not in the climate forecast area.

The recognition that decision-making processes are biased by a variety of problems can lead to improved climate forecast preparation and dissemination. Strategies to reduce the impact of the various cognitive illusions and biases discussed here, in the context of climate prediction, could include some of the following.

1) Ensure groups preparing forecasts undertake debiasing, to reduce overconfidence, especially overconfidence due to hindsight bias or group con-

formity. This would involve forcing forecast groups to actively search for counterexamples (e.g., an example of an El Niño that did *not* coincide with an Australian drought) or misforecasts, or to think of reasons why the forecast may be wrong (e.g., incomplete model).

2) Take care to ensure that media reports and forecasts do not cause anchoring to extreme events such as the 1982/83 El Niño. Every media report noting the seriousness of the impacts for that event could, for instance, also note other El Niño events where the local impacts were less severe.
3) Undertake research to determine how potential users react to probability climate forecasts, including both worded probability statements and numerical values. For instance, determine whether users typically associate the word "likely" with a probability of 70%. Once the relationship between worded and numerical probability forecasts is determined, ensure that media releases and media reports only use terms correctly relating the objective, numerical probabilities.
4) Take care in writing forecasts to ensure that users do not get caught by the way the forecast is framed. This may require the forecast to be expressed in different ways, in the same forecast message. This would be repetitious, but would help users recognize the equivalence of a forecast stated in various ways. This would also help users avoid other sources of bias in decision making, such as decision regret, and asymmetry between loss and gain.
5) Do not prepare forecasts subjectively based on the combination of forecasts or information from various sources. Such an "intuitive" approach is likely to be less than optimal, and objective approaches for combining forecasts and information do exist. Especially, do not combine forecasts within a group, except by simply averaging the forecasts made by individuals or individual models before the group is convened.
6) Ensure that the forecasts are prepared and disseminated in such a way as to avoid users getting caught by ignoring base rates. Thus, base rates would need to be included explicitly when the accuracy of a forecast is presented. The only simple exception would be a forecast of the probability that the predictand would be below (or above) the median.
7) Use external aids (Dawes 1988) when presenting forecasts. Thus, in a forecast of rainfall probabilities dependent on a forecast of El Niño, we might include time series of rainfall and the Southern Oscillation index (to show that the relationship does not always work) and we could also show time series of the "forecasts" of the El Niño (to show that these are also sometimes in error). Such visual aids, rather than or in addition to numerical probability estimates, may help users and forecasters appreciate the real levels of skill.

The relevance of the various paradoxes and biases discussed here to the medical, psychological, legal, and business management fields has received considerable scrutiny. Journals and societies have been established to foster decision making under uncertainty in these fields. Yet there has been very little attention to this in the climate or, more generally, the meteorological field. We attempt to value climate predictions, and to understand how forecast recipients should use forecasts, without consideration of the cognitive basis for decision making under uncertain climate predictions. It is to be hoped that increasing the awareness of forecast providers to these problems will lead to forecasts more likely to be used optimally.

Acknowledgments. My family and colleagues at BMRC, especially those in the Climate Group, provided a great deal of input to this paper, willingly acting as subjects as I tested some of the biases and illusions discussed in this paper. Baruch Fischhoff, Roger Pielke Jr., Tahl Kestin, Mickey Glantz, and two anonymous reviewers kindly provided extensive comments on an earlier version of the manuscript. As always, the librarians of the Bureau of Meteorology provided much help in obtaining papers from many disciplines.

References

Adams, R. M., K. J. Bryant, B. A. McCarl, D. M. Legler, J. O'Brien, A. Solow, and R. Weiher, 1995: Value of improved long-range weather information. *Contemp. Econ. Policy,* **13,** 10–19.

Arkes, H., 1999: Overconfidence in judgmental forecasting. *Principles of Forecasting: A Handbook for Researchers and Practitioners,* J. S. Armstrong, Ed., Kluwer, in press.

Armstrong, J. S., Ed., 1999: Introduction. *Principles of Forecasting: A Handbook for Researchers and Practitioners,* Kluwer, in press.

Asch, S. E., 1946: Forming impressions of personality. *J. Abnorm. Soc. Psychol.,* **41,** 258–290.

Baker, E. J., 1995: Public response to hurricane probability forecasts. *Prof. Geogr.,* **47,** 137–147.

Baron, J., 1994: *Thinking and Deciding.* 2d ed. Cambridge University Press, 588 pp.

Bazerman, M. H., 1994: *Judgement in Managerial Decision Making.* 3d ed. Wiley, 226 pp.

Cerf, C., and V. Navasky, 1998: *The Experts Speak.* Villard Books, 445 pp.

Changnon, S. A., J. M. Changnon, and D. Changnon, 1995: Uses and applications of climate forecasts for power utilities. *Bull. Amer. Meteor. Soc.,* **76,** 711–720.

Dawes, R. M., 1988: *Rational Choice in an Uncertain World.* Harcourt Brace Jovanovich, 346 pp.

——, D. Faust, and P. E. Meehl, 1989: Clinical versus actuarial judgment. *Science,* **243,** 1668–1674.

DeBondt, W., and R. H. Tahler, 1986: Does the stock market overreact? *J. Finance,* **40,** 793–807.

Fischhoff, B., 1975: Hindsight ≠ foresight. The effect of outcome knowledge on judgement under uncertainty. *J. Exp. Psychol.: Human Perception and Performance,* **1,** 288–299.

——, 1994: What forecasts (seem to) mean. *Int. J. Forecasting,* **10,** 387–403.

——, P. Slovic, and S. Lichtenstein, 1977: Knowing with certainty: The appropriateness of extreme confidence. *J. Exp. Psychol.: Human Perception and Performance,* **3,** 552–564.

Gigone, D., and R. Hastie, 1997: Proper analysis of the accuracy of group judgments. *Psychol. Bull.,* **121,** 149–167.

Glantz, M., 1977: The value of a long-range weather forecast for the West African Sahel. *Bull. Amer. Meteor. Soc.,* **58,** 150–158.

Goodie, A. S., and E. Fantino, 1996: Learning to commit or avoid the base-rate error. *Nature,* **380,** 247–249.

Grove, W. M., and P. E. Meehl, 1996: Comparative efficiency of information (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychol. Public Policy, Law,* **2,** 293–323.

Hammer, G. L., D. P. Holzworth, and R. Stone, 1996: The value of skill in seasonal climate forecasting to wheat crop management in a region of high climatic variability. *Aust. J. Agric. Res.,* **47,** 717–737.

Hulme, M., and Coauthors, 1992: Seasonal rainfall forecasting for Africa. Part II—Application and impact assessment. *Int. J. Environ. Stud.,* **40,** 103–121.

Janis, I. L., 1982: *Groupthink: Psychological Studies of Policy Decisions and Failures.* Houghton Mifflin, 351 pp.

Jevons, W. S., 1958: *The Principles of Science.* Dover, 726 pp.; original published 1874 by Macmillian, 943 pp.

Kahneman, D., and A. Tversky, 1982: The psychology of preferences. *Sci. Amer.,* January, 136–142.

Kerr, R. A., 1998: Models win big in forecasting El Niño. *Science,* **280,** 522–523.

Keynes, J. M., 1935: *The General Theory of Employment, Interest and Money.* Harcourt Brace, 403 pp.

Konold, C., 1989: Informal conceptions of probability. *Cognit. Instruction,* **6,** 59–98.

Lamb, P. J., 1981: Do we know what we should be trying to forecast—Climatically? *Bull. Amer. Meteor. Soc.,* **62,** 1000–1001.

Lord, C. G., L. Ross, and M. R. Lepper, 1979: Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *J. Personality Social Psychol.,* **37,** 2098–2109.

Matthews, R. A. J., 1996: Base-rate errors and rain forecasts. *Nature,* **382,** 766.

Mjelde, J. W., T. N. Thompson, C. J. Nixon, and P. J. Lamb, 1997: Utilising a farm-level decision model to help prioritise future climate prediction needs. *Meteor. Appl.,* **4,** 161–170.

Murphy, A. H., 1996: Comments on "Operational omission and misuse of numerical precipitation probability expressions." *Bull. Amer. Meteor. Soc.,* **77,** 1586–1588.

——, and B. G. Brown, 1983a: Forecast terminology: Composition and interpretation of public weather forecast. *Bull. Amer. Meteor. Soc.,* **64,** 13–22.

——, and ——, 1983b: Interpretation of some terms and phrases in public weather forecasts. *Bull. Amer. Meteor. Soc.,* **64,** 1283–1289.

——, S. Lichtenstein, B. Fischhoff, and R. L. Winkler, 1980: Misinterpretations of precipitation probability forecasts. *Bull. Amer. Meteor. Soc.,* **61,** 695–701.

Nicholls, N., 1999: Impediments to the use of climate predictions. *Applications of Seasonal Climate Forecasting in Agricultural and Natural Ecosystems—The Australian Experience,* G. Hammer, N. Nicholls, and C. Mitchell, Eds., Kluwer, in press.

——, and T. Kestin, 1998: Communicating climate. *Climate Change,* **40,** 417–420.

Piattelli-Palmarini, M., 1994: *Inevitable Illusions.* Wiley, 242 pp.

Pielke, R. A., Jr., 1999: Who decides? Forecast and responsibilities in the 1997 Red River flood. *Amer. Behav. Sci. Rev.,* in press.

Plous, S., 1993: *The Psychology of Judgement and Decision Making.* McGraw-Hill, 302 pp.

Pulwarty, R. S., and K. T. Redmond, 1997: Climate and salmon restoration in the Columbia River basin: The role and usability of seasonal forecasts. *Bull. Amer. Meteor. Soc.,* **78,** 381–397.

Redelmeier, D. A., and E. Shafir, 1995: Medical decision making in situations that offer multiple alternatives. *J. Amer. Med. Assoc.,* **273,** 302–305.

Rosenthal, R., and K. L. Fode, 1963: The effect of experimenter bias on the performance of the albino rat. *Behav. Sci.,* **8,** 183–189.

Russo, J. E., and P. J. H. Schoemaker, 1989: *Decision Traps.* Doubleday, 281 pp.

Sink, S. A., 1995: Determining the public's understanding of precipitation forecasts: Results of a survey. *Natl. Wea. Dig.,* **19,** 9–15.

Slovic, P., 1987: Perception of risk. *Science,* **236,** 280–285.

——, and B. Fischoff, 1977: On the psychology of experimental surprises. *J. Exp. Psychol.: Human Perception and Performance,* **3,** 544–551.

Sniezek, J. A., and R. A. Henry, 1989: Accuracy and confidence in group judgement. *Organ. Behav. Hum. Decis. Process.,* **43,** 1–28.

Stewart, T. R., 1997: Forecast value: Descriptive decision studies. *Economic Value of Weather and Climate Forecasts,* R. W. Katz and A. H. Murphy, Eds., Cambridge University Press, 147–181.

——, 1999: Improving reliability of judgmental forecasts. *Principles of Forecasting: A Handbook for Researchers and Practitioners,* J. S. Armstrong, Ed., Kluwer, in press.

——, K. F. Heideman, W. F. Moninger, and P. Reagan-Cirincione, 1992: Effects of improved information on the components of skill in weather forecasting. *Organ. Behav. Hum. Decis. Process.,* **53,** 107–134.

Trenberth, K. E., 1997: Short-term climate variations: Recent accomplishments and issues for future progress. *Bull. Amer. Meteor. Soc.,* **78,** 1081–1096.

Tversky, A., and D. Kahneman, 1974: Judgement under uncertainty: Heuristics and biases. *Science,* **185,** 1124–1131.

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

Vislocky, R. L., J. M. Fritsch, and S. N. DiRienzo, 1995: Operational omission and misuse of numerical precipitation probability expressions. *Bull. Amer. Meteor. Soc.,* **76,** 49–52.

Wason, P. C., 1960: On the failure to eliminate hypotheses in a conceptual task. *Quart. J. Experimental Psychology,* **12,** 129–140.

Weinstein, N. D., 1989: Optimistic biases about personal risks. *Science,* **246,** 1232–1233.