Knowledge Calibration: What Consumers Know and What They Think They Know

JOSEPH W. ALBA
J. WESLEY HUTCHINSON*

Consumer knowledge is seldom complete or errorless. Therefore, the self-assessed validity of knowledge and consequent knowledge calibration (i.e., the correspondence between self-assessed and actual validity) is an important issue for the study of consumer decision making. In this article we describe methods and models used in calibration research. We then review a wide variety of empirical results indicating that high levels of calibration are achieved rarely, moderate levels that include some degree of systematic bias are the norm, and confidence and accuracy are sometimes completely uncorrelated. Finally, we examine the explanations of miscalibration and offer suggestions for future research.

Consumers are overconfident—they think they know more than they actually do. Our simple goal is to evaluate this proposition. Ultimately, we conclude that overconfidence is indeed a robust phenomenon and can be adopted by researchers as a stylized fact about human cognition; however, there are critical qualifications and exceptions that must be kept in mind. The central construct in our analysis is the "calibration of consumer knowledge," which we define as the agreement between objective and subjective assessments of the validity of information—particularly the information used in decision making. For example, a shopper might be very confident that she has found an item at the lowest price available. If she subsequently finds that many other stores offer the same item at a lower price, we would say her knowledge is not well calibrated. Alternatively, if she finds that most other stores have higher prices, we would say she is well calibrated—even if one or two stores have the item at a slightly lower price. Conversely, if she is very uncertain about her price judgments, but is frequently correct, she is also poorly calibrated. That is, calibration refers to the match between confidence and accuracy, rather than accuracy itself.[1]

Calibration is important in consumer decision making because it allows consumers to cope with incomplete and errorful information. Normative theory describes optimal policies for factoring such uncertainties into current decisions and for investing resources to reduce the uncertainty. Clearly, miscalibration is a barrier to implementing such policies. Although much consumer research has examined the extent to which actual decision processes approximate optimal decision policies, there is relatively little research about calibration and miscalibration. Instead, consumer research has focused on knowledge, per se. For example, in a previous paper we attempted to characterize the different facets of consumer expertise (Alba and Hutchinson 1987). In so doing, we distinguished true expertise from mere product familiarity. Whereas familiarity accumulates naturally from any product-related encounter (whether its effects are beneficial or not), expertise is measured relative to a performance criterion and implies increased ability. A related distinction can be made with regard to calibration. Accuracy is dependent on one's ability—or expertise. Confidence is also based on expertise, but other factors, including experience, may influence confidence while leaving accuracy unchanged. Thus, accuracy reflects what we know, confidence reflects what we think we know, and calibration reflects their correspondence.

Consumer research on metaknowledge has largely been restricted to assessment of the correspondence between objective and subjective knowledge (e.g., factual tests vs. self-reported expertise) and has been applied to a relatively narrow range of consumer behaviors (see Park, Mothersbaugh, and Feick 1994). A review of this research

*Joseph W. Alba is Distinguished Professor of Marketing, University of Florida, Gainesville, FL 32611-7155; e-mail: albaj@dale.cba.ufl.edu. J. Wesley Hutchinson is Professor of Marketing, the Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6371; e-mail: wes@marketing.wharton.upenn.edu. The authors wish to thank Jonathan Baron, Jim Bettman, Lisa Bolton, Lyle Brenner, Eric Eisenstein, Abba Krieger, Mary Frances Luce, John Lynch, and the *JCR* reviewers for their helpful comments.

[1]We use the term calibration in its most inclusive sense, covering all types of discrepancies between confidence and accuracy. Several of the research domains reviewed here use the term in more specific and technical senses. Therefore, the reader should bear this in mind when pursuing the cited literature.

yields two interesting observations. First, the level of correspondence between objective and subjective knowledge is not high, indicating some degree of miscalibration. Second, operationalization of the relevant constructs occurs at a relatively abstract level. As we demonstrate throughout the remainder of this article, calibration is a multidimensional construct that can exert effects on decision making in a variety of ways. One clear implication is that consumer research will benefit from adopting (and adapting) the more refined experimental and theoretical approaches used in other domains.

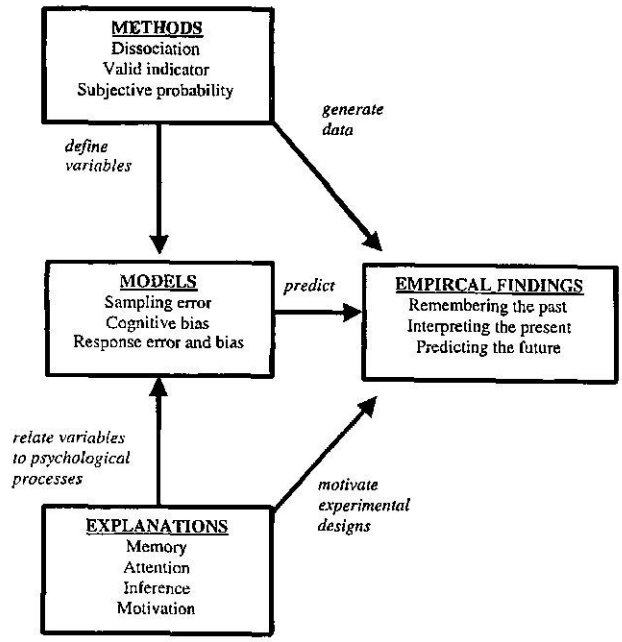
Figure 1 provides an overview of the article and how its component parts are interrelated. We begin with a discussion of the methods and models that have been used in calibration research. Our goal is to develop an abstract account of the experimental paradigms, the data generated by these paradigms, and the sources of miscalibration that have been hypothesized to account for those data. This abstract perspective also provides a bridge between calibration research and traditional problems in consumer behavior. After these preliminary discussions, we review a broad and diverse set of specific research results. These empirical findings can be grouped into three general cognitive activities that produce the information that supports decision making: remembering the past, interpreting the present, and predicting the future. Across these domains the picture that emerges is one in which high levels of calibration are achieved rarely, moderate levels that include some degree of systematic bias (usually overconfidence, but sometimes underconfidence) are the norm, and confidence and accuracy are completely uncorrelated in some situations. Following this exposition of basic findings, we examine the theoretical explanations that have been given for calibration phenomena. These explanations relate the abstract variables of calibration models to specific psychological processes, such as memory, attention, inference, and motivation. Finally, we identify some unanswered questions that are relevant for future consumer research.

I. METHODS AND MODELS IN CALIBRATION RESEARCH

This section provides an abstract overview of the methods and models used in calibration research. We begin by introducing three stylized decision problems to show when calibration will and will not affect decision making and to provide illustrations for subsequent discussions. We then describe the experimental paradigms used in calibration research, the data they generate, and the types of conclusions that can be drawn from such data. In the second part of this section we describe the components of typical calibration models and the sources of miscalibration that are represented in those models. The framework we present is an amalgam of methods and models that have been used across the wide variety of research domains that are reviewed in the remainder of the article. It is rare for a single piece of research to combine all (or even most) of the methods and models that we describe. Thus, we hope to provide both a device for unifying (or at least classifying) disparate perspectives and a source of cross-fertilization for future research.

FIGURE 1

A FRAMEWORK FOR CALIBRATION RESEARCH



A. Calibration and Decision Making: Prototypical Methods

Traditional approaches to decisions under uncertainty have employed choices among well-defined gambles. In this section, we define several simple decision problems of this type and use them to describe the experimental paradigms and measurement procedures used in classic studies of calibration and confidence. We will also use them in the next section to illustrate calibration models and their implications for consumer research.

Three prototypic decision problems that depend on the outcome of a tossed coin are defined in Exhibit 1. Problem 1 is a simple choice task, problem 2 is a simple willingness-to-pay task, and problem 3 combines choice with willingness to pay. In all three tasks the decision maker must form a belief about the likelihood of a heads-up coin toss. In connecting calibration research to consumer decision making it is useful to use these three problems as reference points. In some situations, confidence may reflect metaknowledge that has limited effects on the decision. For example, in a task similar to problem 1, consumers might need to choose which of two restaurants, McDonalds or Burger King, will provide a better meal. Observed behavior may not reflect confidence because a consumer who confidently believes that Mc-

EXHIBIT 1

PROTOTYPICAL DECISION PROBLEMS

Problem 1: Choosing among Uncertain Events
A coin will be tossed. You choose heads or tails. If you choose correctly, you win $100; otherwise, you win nothing. You must decide whether to choose heads or tails.

Problem 2: A Willingness-to-Pay Decision
A coin will be tossed. If it lands heads up, you win $100; otherwise, you win nothing. You must decide the maximum price you are willing to pay to play this gamble.

Problem 3: A Choice-Conditional Willingness-to-Pay Decision
A coin will be tossed. You must choose heads or tails. If you choose correctly, you win $100; otherwise, you win nothing. You must decide the maximum price you are willing to pay to play this gamble.

Donalds is better and one who doubtfully agrees will both go to McDonalds. However, if the high-confidence consumer is sending a friend to McDonalds to get a lunch and says "Don't spend more than $10," we would expect the low-confidence consumer to state a lower spending limit. Thus, problem 2 requires that confidence be factored into the decision. Similarly, if Burger King is located five minutes closer, the high-confidence consumer might stick with McDonalds but the low-confidence consumer might switch. Thus, choice problems like problem 3 require that confidence be used to compare alternatives. More generally, confidence will affect decisions to the extent that consumers use subjective probabilities to evaluate expected values (normative or otherwise), as in problems 2 and 3. Such expected values are commonplace and occur any time two uncertain decision inputs must be integrated. It is only for the simplest of choices, such as problem 1 (in which the decision is determined by a single probability), that confidence is unlikely to affect choice because all levels of confidence above a given cutoff (typically .5) lead to the same decision.

1. Experimental Paradigms

Much of the research that we review presents subjects with a set of problems and observes both the response chosen and self-reported confidence that the response is correct. In the vast majority of cases, the research is on calibration, per se (and not on the effects of calibration on choice), so the problems most resemble problem 1. The coin toss is typically a proposition to be verified as true or false (perhaps a fact about the world or a prediction about future events). Thus, it is a fundamental assumption of our review that the calibration effects observed in these studies are likely to generalize to decisions more like problems 2 and 3.

Calibration research tends to use simple problems because the primary interest is in the extent to which confidence matches accuracy. That is, are people accurate when they are confident and confident when they are accurate? There are three frequently used experimental paradigms that offer increasingly rich sources of data for understanding calibration. All three paradigms use some measure of confidence that is reasonably assumed to be at least ordinally consistent with subjects' beliefs about accuracy, typically a rating scale of some sort.

First, some research uses a dissociation paradigm. In this research, it is not necessary for the experimenter to have an independent and valid indicator of the correctness of the response provided by subjects. Rather, the manipulated experimental factors are known to be unrelated to correctness. If the observed measures of confidence exhibit effects of these manipulations, then some level of dissociation between confidence and accuracy has been demonstrated, and the direction of the effect reveals the direction of the bias (i.e., more confident or less confident, but not the absolute level of over- or underconfidence). For example, if familiarity with a product claim is manipulated experimentally (say through number of advertising exposures) and that manipulation increases confidence, then a dissociation has been demonstrated because the familiarity manipulation cannot logically affect the validity of the claim.

Second, some research uses a valid-indicator paradigm. Research of this type requires the experimenter to have an independent and valid indicator of the correctness of the response provided by subjects. Unlike the dissociation paradigm that holds response correctness constant, problems in this paradigm are chosen such that variation in response correctness is likely. For example, we might sample consumers who have recently joined a fitness program and measure their confidence that they will exercise on a regular basis and their actual workout frequency. The extent to which confidence is itself a valid indicator of correctness (e.g., correlations between confidence and workout frequency) provides a measure of calibration. In this paradigm, it is not possible to directly assess over- and underconfidence, per se, because confidence and accuracy are measured on different scales.[2]

Finally, some research uses a subjective probability paradigm. This is the classic approach that introduced the term "calibration" into the decision-making literature (see Adams and Adams 1961; Lichtenstein, Fischhoff, and Phillips 1982). In this paradigm, subjects are carefully instructed to use the response scale such that the number chosen corresponds to their belief about the probability that the response they chose is the correct response. Although the exact procedures required to obtain such a measure can be debated, the concept is straightforward. In addition to the correlation between confidence and accuracy, it is possible to examine over- and underconfidence by direct comparison of subjective and objective probabilities. Although a number of measures of calibration have been used, the most robust and frequently reported measure is the simple difference between the average stated probability and the average percent correct (which is called the "mean over/underconfidence score"). In the next section, we describe this paradigm in some detail because it provides a useful conceptual structure

[2]See Clark (1997) for an example of how somewhat stronger conclusions can be drawn when dissociation and valid-indicator paradigms are combined.

EXHIBIT 2

OBJECTIVE DECISION RULE

An Objective Decision Rule: The Sample-Based Average as a Criterion
Prior to the toss that will determine the outcome of the gamble, experimentally toss the coin n times, observing h heads and t tails. Define $\hat{p} = h / n$. If $\hat{p} > .5$ then choose heads in problems 1 and 3; otherwise, choose tails. In problem 2, the maximum price is $\hat{p} \times \$100$. In problem 3, the maximum price is $\hat{p} \times \$100$ if heads was chosen, and it is $(1 - \hat{p}) \times \$100$ if tails was chosen.

for understanding how calibration, and miscalibration, can affect decision making.

2. Classic Calibration Studies

In a traditional calibration experiment, subjects are asked to state a belief and then assess the likelihood that the belief is valid. In the context of coin tosses, the belief is whether the coin will land heads up or not. Thus, in a calibration study using problem 1, the decision maker would choose heads or tails (i.e., state her belief about the outcome of the coin toss) and then subjectively estimate the probability that this outcome will occur. This estimated probability will range between .5 and 1 because had the decision maker believed his or her choice to have a probability less than .5, the alternative choice would have been made. This subjective probability task is called a "half-range" task. In problem 2, the decision maker would state a maximum price and then report the subjectively estimated probability that the coin will land heads up. This probability can take on any value between zero and one and is called a "full-range" task.

Confidence is conceptualized as degree of certainty. It is highest when subjective probabilities are extreme (i.e., close to either zero or one) and lowest when they are equivalent to random guessing (i.e., .5 when there are two possible outcomes, $1/k$ when there are k outcomes). Therefore, reported subjective probabilities from a full-range task are typically converted to half-range equivalents. This is both a technical and a conceptual aspect of the paradigm. It results from defining confidence as degree of certainty. It is different from subjective assessments of likelihood, per se. A person might overestimate the probability of beating Michael Jordan in one-on-one basketball as 1 out of 100 rather than, say, 1 out of 1,000,000. From the perspective of calibration research, however, this reflects underconfidence in the belief that a loss will occur. It places more uncertainty in the outcome than actually exists. Thus, over/underconfidence scores are computed as if the task were a half-range task.[3] Positive scores indicate overconfidence, and negative scores indicate underconfidence.

Calibration is highest when the over/underconfidence scores for all problems are close to zero; however, calibration can be low even when the mean over/underconfidence score is near zero if the individual scores are large in absolute magnitude (but cancel out across problems). Therefore, the mean squared over/underconfidence score is often computed.[4] In connecting classic calibration research to that using other paradigms it is important and useful to note that dissociation paradigms provide evidence about relative levels of over- or underconfidence in addition to levels of calibration (insofar as they reveal the direction of bias caused by some factor that affects confidence but not accuracy). Valid-indicator paradigms mainly assess level of calibration (because most correlation measures are unaffected by differences in the intercepts, i.e., over- or underconfidence).

B. Calibration and Decision Making: Prototypical Models

Models of calibration and decision making must specify (1) the information used to make the focal decision and the confidence judgment for that decision, (2) the decision rule that is applied to that information for the decision and the judgment, and (3) the temporal flow of information, decisions, and judgments (possibly including covert decisions and judgments that precede observed behaviors). In this section we sketch out several types of models that are found in calibration research.

1. Decision Rules

The prototypical decision rule for calibration tasks is that a covert subjective probability is computed, and this probability is used both to make a specific decision and assess one's confidence in that decision. Exhibit 2 describes an

[3]For example, in problem 2 (a full range task), reported probabilities greater than .5 are assumed to indicate a prediction of heads (with a subjective probability equal to the reported probability), and reported probabilities less than .5 are assumed to indicate a prediction of tails (with subjective probability of 1 minus the reported probability for heads).

[4]This is called a Brier score and it is often decomposed into a calibration component, a resolution component, and a knowledge component (which is unrelated to judgment). Although we use the term calibration in a more general sense, the calibration score is more or less consistent with that usage. Resolution reflects the extent to which different confidence levels are associated with different levels of accuracy. However, resolution does not require that the association between confidence and accuracy be monotonic. Thus, resolution is seldom used in recent research, and various measures of correlation are generally preferred. Maki (1998b) provides an excellent discussion of the measures of metacognitive accuracy used in different research domains. For more detailed discussion of measurement issues, see Brenner (1999), Erev, Wallsten, and Budescu (1994), Klayman et al. (1999), and Liberman and Tversky (1993).

"objective" decision rule for the three prototypical problems described earlier. This rule uses a sample-based average as the choice criterion. Before the coin is tossed, the decision maker is allowed to acquire uncertainty-reducing information about the coin by making some number of preliminary tosses of the same coin that is to be used in the gamble. The decision maker then computes the sample mean (using heads = 1, tails = 0). We call this an objective rule because the sample mean is an unbiased estimator of the probability that the coin will land heads up in the critical toss that determines the payoff of the gamble. This provides a convenient benchmark for psychological models because we can easily replace the preliminary sampling with other uncertainty-reducing processes, including the perception and comprehension of explicitly provided information and internally generated information based on memory or inferential reasoning.

The sample-based average criterion maximizes expected winnings in problem 1. As discussed earlier, the probability of heads, per se, is not important. The only critical information coming from the sample of coin tosses is whether there were more heads than tails, or vice versa. Problem 1 provides an example of where calibration problems will have limited effects on decision making because only when objective and subjective probabilities fall on opposite sides of .5 will the two lead to different decisions. Such effects are likely only when there are very large biases or when problems have high uncertainty (i.e., probabilities near .5).

In contrast, problems 2 and 3 require the decision maker to use metric properties of the subjectively estimated probability. For example, when the payoff is held constant, it is safe to assume that most people would be willing to pay more when their subjective probability of winning is higher. Calibration is therefore a more central component of decision making. The sensitivity of the decision to calibration is important because it should affect the willingness of the decision maker to devote resources to reducing uncertainty (i.e., to building confidence in their knowledge).

For example, assume that a decision maker faced with problem 1 observed four heads in five tosses of the coin prior to making a choice. This person is then offered the opportunity to observe one (and only one) more toss at a cost of $1. If the person is using the sample-based average as a decision criterion, this offer should be declined because it will result in a sample average of either .67 (4 out of 6) or .83 (5 out of 6), and neither outcome will change the decision based on the current sample average of .80 (4 out of 5). However, suppose the price of playing the gamble is $80. Now the decision maker is faced with problem 3 and in using the sample-based average is indifferent between playing the gamble and not playing because the expected value of the gamble is exactly $80 (i.e., .80 × $100). Is a sixth toss of the coin worth $1 to this person? On the basis of the current sample average, this person believes there is a 20 percent chance that a tail will be observed and an 80 percent chance that a head will be observed. If a tail is observed the gamble will be valued at −$13 (i.e., $67 − $80) and will therefore be declined (i.e., a certain outcome of $0). If a head is observed, the gamble will be valued at $3 (i.e., $83 − $80) and will be accepted. Thus, the expected value of a sixth coin toss is $2.40 (i.e., .2 × $0 + .8 × $3) and is therefore worth more than its $1 cost.

We learn two things from this example. First, we learn that even apparently simple decisions are greatly complicated by the opportunity to reduce uncertainty. This is because one must accurately assess the value of future information and options. For such problems, both optimal solutions and actual decision processes are necessarily complex (Hutchinson and Meyer 1994). Second, and more important, we learn that the effects of calibration (and miscalibration) depend on the specific decision. Calibration will have limited effects on decisions of pure prediction but can be centrally important when a value must be placed on an uncertain outcome or on activities that will reduce uncertainty. We will return to this issue at the conclusion of this section.

2. Decision Processes and Sources of Miscalibration

Recent calibration research using the subjective probability paradigm has convincingly demonstrated that miscalibration can arise from a variety of sources, and some of those sources are rather different from what people have assumed in the past. In addition to the various cognitive heuristics and biases that have been proposed, the presence of random error in the decision process can generate biases in the subjective probabilities reported by subjects even when the error, per se, is unbiased (discussed in more detail in the Empirical Findings section). In certain situations, these stochastic sources are best regarded as due to experimental design or measurement and do not imply that decision making is biased as the result of miscalibration. In other situations, random error is an integral part of underlying cognitive processes (e.g., perception, memory, attention), so miscalibration is likely to lead to biased decisions outside of the laboratory in real consumer decisions. Thus, it is important to identify the sources of miscalibration. Unfortunately, in the very large subjective probability literature and in the even larger literature on metaknowledge, very few studies have been designed specifically to assess the relative contributions of different sources of miscalibration. In the review that follows, we will report evidence about sources whenever possible. However, in the majority of cases, experimental evidence will reveal the existence of miscalibration but not its specific source. This is obviously a fertile area for future research. In this section we provide a concise overview of the sources of miscalibration that have been discussed most often in the literature.

Consider again the decision maker who must decide whether to pay $80 to gamble on the outcome of a coin toss that will pay $100 if she is correct and nothing otherwise. In addition to deciding on the gamble, the decision maker

must also report her subjective probability of winning.[5] Exhibit 3 describes two possible decision processes that could be used in this task. In both processes, the decision maker first samples information. This information is a series of tosses of a test coin that is similar to the coin used in the gamble. From a more general perspective, this step represents whatever informational inputs are available to the decision maker (e.g., based on the perception and comprehension of externally available information or on internally generated information resulting from memory or inference). The test coin has a probability of coming up heads, p, that may or may not be the same as that of the gamble coin, p^*. In process 1, the decision maker observes $\hat{p}$, uses this value to decide whether or not to gamble, and then reports a subjective probability, $r(\hat{p})$, that may or may not be the same as $\hat{p}$. Error and bias could enter the decision process at any point. However, any error or bias that enters at step 3 will affect calibration but it will not affect the gambling decision. In contrast, process 2 reports confidence prior to the gambling decision and uses that reported value to make the decision. Thus, miscalibration at any stage affects the decision.

The key difference between the two processes is not the order of the steps, per se, but whether the sources of miscalibration that affect reported confidence also affect decision making. Generally speaking, "sampling error" refers to unbiased random error that occurs in step 1 of process 1 (and process 2).[6] For example, even though the long-run expected value of $\hat{p}$ is p, on any given occasion p and $\hat{p}$ will differ by some amount. "Cognitive biases" refer to systematic tendencies to sample the wrong types of information (e.g., p differs from p^* in step 1) or to decision rules that are not optimal (in step 2). "Response error" refers to unbiased random error that occurs in step 3 of process 1 (and step 2 of process 2). For example, even if the long-run expected value of $r(\hat{p})$ is $\hat{p}$, on any given occasion $r(\hat{p})$ and $\hat{p}$ will differ by some amount when response error is present. "Response biases" refer to systematic tendencies for the long-run expected value of $r(\hat{p})$ to differ from $\hat{p}$.

Because most of the results that we describe do not isolate the sources of miscalibration, consumer researchers must apply them cautiously. Some decision processes are plausibly assumed to resemble process 2 in Exhibit 3. For example, in group decision making, personal selling, and word-of-mouth communications, verbally expressed confidence (i.e., $r(\hat{p})$) is used directly in subsequent decisions and judgments. Similarly, extended decision making by an individual may be assumed to consist of an information-gathering phase during which confidence is explicitly assessed and remembered prior to a final decision (even if it was never verbalized). In these situations, all sources of miscalibration are likely to affect consumer decisions. However, some situations may be more like process 1, in which response error and bias affect judged confidence but not the focal decision. Such situations require methods designed to identify the specific source of miscalibration, or at least rule out mere response effects. For example, experiments testing the hypothesis that advertising increases consumer overconfidence in ways that favor the advertised products should be designed to rule out response effects because the hypothesis is about biases that will affect consumer choice. More generally, research that focuses on decision processes more than decision outcomes needs to examine the sources of miscalibration in addition to observing effects on measures of confidence and calibration.

EXHIBIT 3

POSSIBLE DECISION PROCESSES

Process 1

Step 1: INFORMATION SAMPLING
Observe a series of tosses of the test coin, and compute $\hat{p}$.

Step 2: APPLY DECISION RULE
If $\hat{p} \times \$100 > \$80$, then bet on heads. If $(1 - \hat{p}) \times \$100 > \$80$, then bet on tails. If neither holds, then decline the gamble.

Step 3: REPORT CONFIDENCE
If $\hat{p} > .5$, then report $r(\hat{p})$. If $\hat{p} < .5$, then report $1 - r(\hat{p})$.

Process 2

Step 1: INFORMATION SAMPLING
Observe a series of tosses of the test coin, and compute $\hat{p}$.

Step 2: REPORT CONFIDENCE
If $\hat{p} > .5$, then report $r(\hat{p})$. If $\hat{p} < .5$, then report $1 - r(\hat{p})$.

Step 3: APPLY DECISION RULE
If $r(\hat{p}) \times \$100 > \$80$, then bet on heads. If $(1 - r(\hat{p})) \times \$100 > \$80$, then bet on tails. If neither holds, then decline the gamble.

In summary, cognitive biases and sampling error degrade the quality of decision making of almost all decisions. Random error and biases that arise during response selection will affect only those decisions that use the response directly as input. Thus, for any specific consumer application, researchers should carefully consider which sources of miscalibration are most important for the problems being addressed.

3. Implications for Consumer Research

Because most of the research that we subsequently review does not explicitly examine the effects of knowledge calibration on consumer behavior, it is useful at this point to say why we believe it is very likely that such effects exist and are important. Of course, it is our hope in writing this review that future research will confirm these expectations.

We illustrate the relationship between calibration and consumer behavior with the following vignette. Ms. Smith has a medical condition that is treatable, but not curable, with drug therapies. The condition spontaneously resolves itself in some people, and Ms. Smith feels she has been getting better. Her doctor tells her about a relatively new surgical procedure that has been found to cure some patients and suggests that she consider this option. The side effects of

[5]The reasoning here also applies to confidence ratings but using subjective probabilities simplifies the discussion.

[6]We are using the term "sampling error" in a fairly inclusive sense, covering all forms of stochastic error in the informational inputs. In our example, this error is described by the binomial distribution.

the drugs she is taking are significant, but bearable. Initially, Ms. Smith decides not to consult a surgeon because she believes she will be one of the lucky ones who recover spontaneously. After a few years of intermittent improvement and decline she decides to visit a surgeon, Dr. Jones. Jones tells her that on the basis of her age and medical history he feels there is a 90 percent chance that the surgery will succeed. The surgery will require a week in the hospital and two or three months for full recovery. Of course, there is a very small chance of unexpected complications with any major surgery. All things considered, Dr. Jones recommends the procedure. Ms. Smith agonizes over the decision and finally agrees to have the surgery.

If we assume that both Smith and Jones are overconfident in all of their beliefs, what are the implications? Overconfidence in the likelihood of spontaneous recovery may have delayed Ms. Smith's information search. Overconfidence may have caused Dr. Jones to state a 90 percent chance of success when the real chance was only 70 percent. Assume that Dr. Jones's overconfidence was a response effect. His decision to recommend surgery may have been unbiased, but Ms. Smith has only the verbal report of 90 percent. Given that the decision was difficult for her, she may have preferred her chances for spontaneous recovery if Jones had said 70 percent. More generally, it seems reasonable that her final decision was affected by her confidence in all of the decision inputs (i.e., spontaneous recovery, successful surgery, avoiding unexpected complications, and length of the recovery period). We also see that some of these inputs are based on past experience, some on interpreting current information, and some on predicting future events.

This vignette illustrates several major ways in which calibration can affect consumer decision making. First, it affects information search. Because the goal of information search is usually to reduce uncertainty, overconfidence will tend to inhibit search and underconfidence will increase search.[7] Second, in many consumer situations both the overt and covert judgments of subjective probability are used directly as inputs to decision making. Third, calibration is likely to affect how consumers resolve any conflict that arises between two or more uncertain decision inputs. Finally, most decisions are based on specific information items; therefore, individual decisions will be affected by poor calibration even when over- and underconfidence tends to cancel out across items. That is, decisions depend on which beliefs are overconfident and which are underconfident—not the average level of over/underconfidence. Given all of the ways that calibration can potentially affect decision making, there is a clear need for consumer-oriented research that specifically applies the findings and explanations discussed in the remainder of this article.[8]

[7]Although this appears to be a robust general trend, it is possible to construct counterexamples where as overconfidence increases there is initially a tendency to search more than is optimal followed by a tendency to search less than is optimal as overconfidence becomes extreme (John Lynch, personal communication).

II. EMPIRICAL FINDINGS

In this section we recount outcomes from a large amount of research on metaknowledge. The extent to which these outcomes are firmly supported versus merely suggested varies widely across phenomena. Consistent with the objectives of this article, we favor comprehensiveness; however, we also note the degree of empirical support for any given phenomenon. Temporal characteristics of the paradigms employed in this research serve as a convenient and intuitive organizational device. In some situations, consumer decisions are based on remembered information, and therefore calibration of memory for past events or previously accumulated knowledge is central. In other situations, consumers must assess the appropriateness of currently held beliefs. In yet other situations, consumers must forecast the likelihood of possible outcomes. The appropriateness of the confidence consumers place in their knowledge, beliefs, and predictions will influence the appropriateness of the actions they take.

A. Remembering the Past

An obviously important source of informational inputs to decision making is memory (Alba, Hutchinson, and Lynch 1991). We review research that examines metaknowledge about remembered information and how confidence in remembered information changes over time. First, we examine memory for factual knowledge, such as "Mercedes are expensive cars." Then, we examine memory for specific events, such as "the salesperson said the monthly payments for a Mercedes 190 would be $400." This distinction between facts and personally experienced events reflects differences in experimental paradigms examining semantic and episodic memory (e.g., Tulving 1983). Finally, we review unconscious learning, which is a source of low-confidence information, almost by definition.

1. Memory for Statements of Fact

The most direct assessment of knowledge calibration is simply to ask an individual to estimate the probability that a recalled piece of information is valid. In fact, there is an extensive literature on exactly this task that uses the subjective probability paradigm. A set of general knowledge questions with unambiguous answers is first assembled (e.g., the populations of cities, the dates of historical events, the prices of consumer products). For each question, subjects either report the probability that a given proposition is true (i.e., a full-range task) or decide whether the proposition is true or false and report the probability that their answer is correct (i.e., a half-range task). This task is not far removed

[8]For an excellent discussion of how lack of calibration (and other cognitive biases) can affect managerial decision making, see Russo and Shoemaker (1989).

from consumer decision making insofar as consumers often must verify a set of facts and claims about choice alternatives as part of their decision-making process.

From the perspective of empirical outcomes associated with this paradigm, several robust findings have been known for a long time. These findings can be summarized as follows: (1) moderate correlations between subjective probabilities (i.e., reported confidence) and objective probabilities (i.e., percent correct); (2) a general tendency toward overconfidence (i.e., the overall mean subjective probability is greater than the overall mean objective probability); (3) when experimental conditions differ in difficulty (e.g., as measured by the average percent correct) there is more overconfidence in the hard than the easy conditions (called the "hard-easy effect"); and (4) within an experimental condition, calibration curves plotting mean percent correct for each confidence level exhibit a characteristic shape in which overconfidence is greatest when confidence is high and least when confidence is low (often reversing to underconfidence). Excellent reviews of these results can be found in Keren (1991), Lichtenstein et al. (1982), and McClelland and Bolger (1994, which focuses on models more than findings).

More specifically, confidence about answers given to general knowledge questions is almost always found to be correlated with percent correct. Typically, these correlations at the individual level are rather low but positive (e.g., $r = .20$ for the average subject; Griffin and Varey 1996). However, high correlations are sometimes observed (e.g., $r = .75$; Erev et al. 1994; see also Klayman et al. 1999 and Maki 1998a). The importance of this simple result is that consumers are better off using confidence as an input to decision making than ignoring it—if they use it appropriately. That is, confidence is generally a valid indicator of the accuracy of knowledge. However, the appropriate-use caveat is critical, and most research in this area can be interpreted as investigating the extent to which people are able to do so.

We should note that there are cases in which researchers using somewhat different paradigms have found different patterns of confidence. Hutchinson and Alba (1991) collected confidence ratings in a consumer learning task. Some subjects did not fully learn the target concept and included irrelevant as well as relevant information in their decision rules. These subjects expressed confidence ratings that—consistent with their rule—initially declined as accuracy declined but increased again when accuracy was lowest. This reversal of the confidence/accuracy relationship should be expected whenever a decision rule uses much more irrelevant information than relevant information. When all irrelevant information points to the wrong answer, it dominates the relevant information and leads to confident judgments that are consistently wrong; when irrelevant information is mixed, confidence will be low but judgments will be more likely to be correct (because relevant information can exert an effect). A similar effect has been observed in the area of memory for personal events (discussed in the next section). The presence of misinformation can actually produce negative correlations between confidence and accuracy. Weingardt, Leonesio, and Loftus (1994) found that subjects given misleading information expressed higher levels of confidence when they were incorrect than when they were correct. Unlike typical calibration experiments, both of these studies had an element of deception insofar as misleading information was made salient to subjects.

A third study that is an exception to the rule (i.e., that confidence is generally a valid indicator for general knowledge problems) is provided by Radecki and Jaccard (1995). Across two important consumer-relevant domains (birth control and nutrition), they found that self-assessed knowledge and actual knowledge were weakly related and, moreover, that search was inversely related to self-assessed knowledge (see Moorman 1999). The correlational nature of this analysis prohibits specific conclusions about over- and underconfidence. Nonetheless, insofar as information search is terminated on the basis of an uncalibrated assessment of knowledge, any suboptimal decisions that result will go undetected. These consumers do not know that they do not know and therefore are unaware of the need to search for more information or take other corrective actions.

Overconfidence. Perhaps the most robust finding in the calibration literature is that, unless the problems are very easy, the overall mean subjective probability is greater than the overall percent correct, indicating overconfidence. The levels of overconfidence that are typically observed range as high as 15 percent. Overconfidence is even more extreme when a confidence-range task is used. In this task, subjects are instructed to state upper and lower bounds on some dimension (e.g., the price of a product) such that the right answer will fall in that range 90 percent of the time. Instead of the requested 90 percent, the range includes the correct answer 40–60 percent of the time (Klayman et al. 1999; Lichtenstein et al. 1982; Russo and Schoemaker 1992).

Clearly, overconfidence is a ubiquitous result. If all problems are rather easy, however, underconfidence may be observed. More generally, experiments manipulating problem difficulty routinely find that reductions in problem difficulty lead to reductions in overconfidence. However, equally robust is the following apparently paradoxical result. When problems are grouped according to confidence level, the greatest overconfidence is observed for the problems answered with the greatest confidence. For example, subjects typically express 100 percent confidence in a substantial number of answers. Accuracy for these problems is typically 80 percent or less. Conversely, when subjects claim to be guessing (i.e., 50 percent confidence levels), accuracy is significantly above chance.

All of these effects are illustrated with simulated data in Figure 2. This figure shows the mean percent correct for each of six levels of rated confidence. Points to the left of the 45-degree line indicate underconfidence and points to the right indicate overconfidence. For the hard set of problems, most of the points are to the right of the line. By itself, this does not imply that the overall average is overconfident because all confidence levels are not used equally often. In

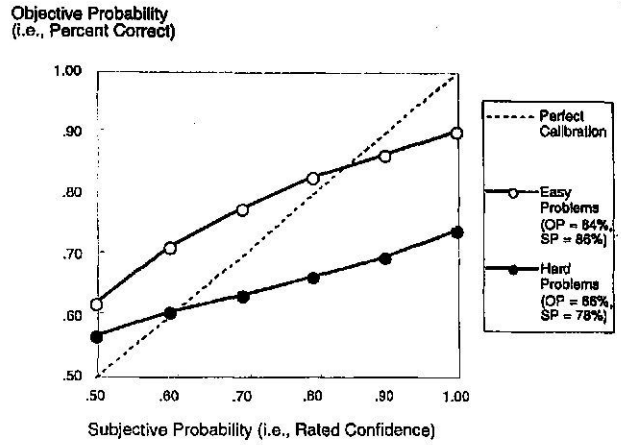
this case, and as is typically true, average confidence (SP) exceeds overall percent correct (OP; i.e., 78 percent > 66 percent). For the easy set of problems, most of the points are to the left of the line. In this case, overconfidence is reduced but not reversed (i.e., SP = 86 percent and OP = 84 percent). In both cases, accuracy is considerably less than perfect when complete certainty is claimed, and accuracy is better than chance when complete uncertainty is claimed.

The pattern of results shown in Figure 2 is extremely robust; however, the interpretation of these results has become very controversial in recent years. In particular, debate has centered on the extent to which these results are due to cognitive biases, sampling error, or response error.[9] A number of simulation studies have convincingly demonstrated that models containing only unbiased sampling and response error can generate most or all of the classic results (i.e., overconfidence, hard-easy effects, and typical calibration curves). This finding is counterintuitive because unbiased error leads to biased responses.[10] Of course, demonstrating consistency with classic results does not prove that these results are in actuality due to stochastic components and not cognitive biases. Moreover, separating the relative contributions of these components is likely to be an arduous task. In the simulation used for Figure 2 (which is based on those of Juslin, Olsson, and Bjorkman [1997]), all three components have been included. To illustrate the complexity with which these sources interact, we note that in the easy condition the mean confidence bias is +2 percent. When simulations were run for each component acting in isolation the resulting biases were +6 percent, +3 percent, and −4 percent for the cognitive, sampling, and response components, respectively.[11] Thus, the net bias is not a simple sum or average of the component biases.

We know of no studies that have experimentally manipulated the various sources of miscalibration independently.[12] However, several studies have estimated statistical models with only random error components (Budescu et al. 1997; Juslin et al. 1997; Soll 1996; Suantak et al. 1996). All four of these studies show that simple stochastic models can fit calibration data well. However, two of the four (i.e., Budescu et al. 1997; Soll 1996) also show that there is at least some systematic residual bias for at least some subjects that cannot be accounted for by the stochastic model and is presumably due to biased cognitive processes. Dougherty, Gettys, and Ogden (1999) adapt Hintzman's (1988) exemplar-based model of memory to decisions made under uncertainty. Using simulation methods, the model is shown to be consistent with a very large number of results in the decision-making literature, including calibration effects discussed here (and the improvements in calibration that accrue with expertise, discussed at the end of this article). The model includes all of the sources of miscalibration and shows how they arise naturally from memory processes. Our tentative conclusion from these studies (and one that is shared by most of these researchers) is that the basic finding of robust overconfidence is likely to result from both random error and cognitive biases. In fact, this is probably why it is such a robust finding.

FIGURE 2

SIMULATED CALIBRATION CURVES



The Hard-Easy Effect. The "paradox" of the hard-easy effect and calibration curves—that is, maximum overconfidence when confidence is high and maximum underconfidence when problems are easy (and confidence is therefore high on average)—is almost certainly due in large part to statistical regression to the mean. Although some have characterized this as a measurement artifact (Erev et al. 1994),

[9]This debate can be found conveniently in several places. The March 1996 issue of *Organizational Behavior and Human Decision Processes* has key articles by Dawes and Mulford; by Brenner et al.; and commentary by Griffin and Varey and by Wallsten. The September 1997 issue of the *Journal of Behavioral Decision Making* has key articles by Budescu, Erev, and Wallsten; by Budescu, Wallsten, and Au; and by Juslin, Olsson, and Bjorkman; and commentary by Ayton and McClelland and by Keren. Also see Klayman et al. (1999), and Suantak, Bolger, and Ferrell (1996).

[10]To see how this can happen, consider our stylized coin toss problem and the objective decision rule of Exhibit 2. Assume that $p = .6$ and three preliminary tosses are permitted. If there are two or three heads then "heads" will be predicted and the expected percent correct is 60 percent; if there are zero or one heads then "tails" will be predicted and the expected percent correct is 40 percent. Thus, the expected overall percent correct is between 40 percent and 60 percent (specifically, it is 53 percent). Now, if zero or three heads are observed, then the reported subjective probability will be 100 percent; if one or two heads are observed, then the reported subjective probability will be 67 percent. Thus, the expected overall subjective probability is between 67 percent and 100 percent (specifically, it is 76 percent). Overconfidence, even for this apparently "objective" decision rule, results from the fact that conditioning on the prediction leads to subjective probabilities that are always higher than the best possible percent correct (i.e., 60 percent).

[11]Similar results obtain for the hard condition where the component biases were +9, +7, and +2 percent and the bias in the combined simulation was +12 percent. Details of the simulation are available from the second author upon request.

[12]Wallsten, Bender, and Li (1999) report experiments on pure verification responses, without explicit confidence judgments, and use related modeling to isolate the effects of sampling error, response error, and response bias.

we agree with those who describe it as a mathematical fact of life (Dawes and Mulford 1996; Griffin and Varey 1996). This paradox results from the conditional nature of the key concepts: problem difficulty and judged confidence. This is most easily seen at the extremes. When problems are maximally easy (i.e., conditioning on problem difficulty), accuracy is 100 percent and any type of error or bias can only make judged confidence less than 100 percent (i.e., underconfidence). Conversely, when confidence is maximally high (i.e., conditioning on judged confidence), reported subjective probabilities are 100 percent and any type of error or bias can only make accuracy less than 100 percent (i.e., overconfidence). The same reasoning leads to overconfidence on hard problems and better-than-chance performance when confidence is minimal. The conclusion is that whenever conditional information, such as confidence, is used as a decision input, regression-to-the-mean biases will result.

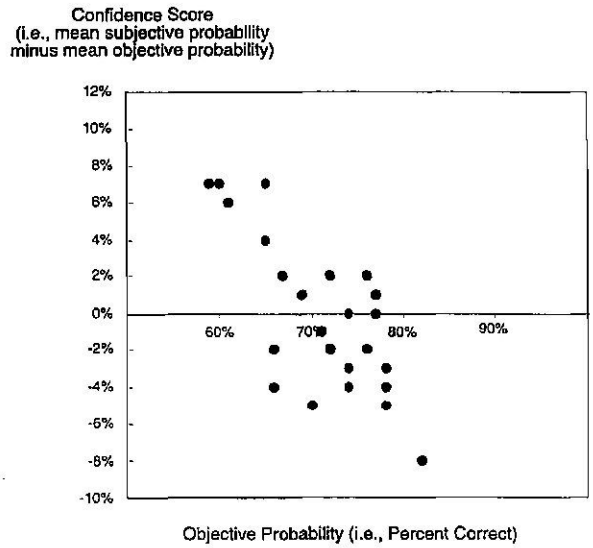
Although regression to the mean provides an explanation of the paradoxical relationship between calibration curves and the hard-easy effect, the hard-easy effect itself has been attacked as a methodological artifact that results from the use of trick questions (e.g., Gigerenzer, Hoffrage, and Kleinbolting 1991). In particular, it is argued that cue-based reasoning will result in miscalibrated confidence when questions are selectively sampled by researchers because the questions artificially change the probabilities that the cues are valid. From a purely empirical standpoint, there is no doubt that question selection can affect calibration, and this reinforces our earlier observation about the power of misinformation and outright deception (e.g., Weingardt et al. 1994). Also, from a purely empirical standpoint, it is now clear that overconfidence and hard-easy effects are often found even when experimenters are very careful about question sampling. Figure 3 shows the results of 25 studies conducted by Juslin and his colleagues (summarized in Juslin et al. [1997]). Although the sizes of the effects are somewhat smaller than have been observed in nonrepresentatively sampled experiments, miscalibration is evident and there is a clear hard-easy effect across experiments.

It should be noted that Klayman et al. (1999) found that there were reliable differences in overconfidence across domains (e.g., calories in different foods vs. tuition for different colleges) and reliable individual differences in overconfidence, but that degree of overconfidence was unrelated to difficulty. Their method removed possible regression-to-the-mean effects by randomly splitting questions in a domain and using one set to assess difficulty and another to assess confidence. Thus, observed hard-easy effects may be mainly due to statistical regression. However, it is important to note that such effects are still important from the perspective of decision making because the accuracy of and confidence in specific information items, not overall averages, determines specific decisions.

Summary. Given the several controversies that we have described, what should we conclude about the classic results and the recent debate? As other authors have pointed out (e.g., Dougherty et al. 1999; Griffin and Varey 1996; Suantak et al. 1996), and as our subsequent review will demonstrate, there is ample evidence from other domains and paradigms that confidence is affected by various cognitive biases. There is also increasing evidence that random error alone cannot explain all of the observed results about the calibration of general knowledge. Perhaps the most important contribution of this debate has been to add mere random error to the list of potentially powerful sources of miscalibration and to show clearly that a more sophisticated approach to modeling metaknowledge is needed. As we noted in an earlier section, all of these sources of miscalibration present a real threat to at least some types of consumer decision making.

FIGURE 3

THE HARD-EASY EFFECT FOR 25 EXPERIMENTAL STUDIES



2. Memory for Events

The most intensively investigated episodic context has involved eyewitness testimony. Calibration is central because, both psychologically and legally, the persuasiveness of a witness's testimony is a function of the confidence with which it is offered. Meta-analyses of the correlation between accuracy and confidence in eyewitness testimony produce results that vary widely (Bothwell, Deffenbacher, and Brigham 1987; Sporer et al. 1995; Wells and Murray 1984). Calibration may be particularly poor when the witness makes a deliberate (rather than automatic) eyewitness recognition (Dunning and Stern 1994). This is counterintuitive because it is natural to think that more deliberation should improve decision making. Recently, the appropriateness of simple correlational indices of calibration has been questioned (Juslin, Olsson, and Winman 1996). Nonetheless, even studies that use more traditional calibration techniques

have provided some indication of significant overconfidence (Olsson, Juslin, and Winman 1998), particularly when misinformation has been provided to the witness (Weingardt et al. 1994; see also Loftus et al. 1989).

Outside the legal context, the study of memory miscalibration is more fragmented. A variety of studies suggest overconfidence, but additional research is clearly in order. Wagenaar (1988) provides results that are consistent with the eyewitness literature but use a range of traditional memory stimuli—including lists, events, and scenarios—and traditional calibration measures. In extreme cases, he finds high confidence in memories that failed to achieve chance levels of accuracy. The effects were strongest when accuracy was low but the particular response seemed plausible. More recent research has shown that false memories may be held with very high confidence and may even be accompanied by conscious feelings of episodic experience—particularly when the falsely remembered information is semantically related to the truly encountered information (Holmes, Waters, and Rajaram 1998; Roediger and McDermott 1995).

In contrast, Schneider and Laurion (1993) report surprisingly high levels of metamemory for incidentally learned news reports (see also Pritchard and Keenan 1999). Subjects in their study listened to a string of news reports and then answered objective questions based on the reports. For each answer, subjects also rated their confidence in the correctness of the answer. The overall confidence-accuracy correlation exceeded .70. As Schneider and Laurion (1993) note, the outstanding performance they obtained may be attributable in part to the ease of the task. Their task required no reasoning or inference but rather immediate memory.

Finally, a handful of studies suggest miscalibration in autobiographical memory. Using a diary method, Barclay and Wellman (1986) found that a decline in accuracy of autobiographical memory over time was not accompanied by a decline in confidence, implying that overconfidence increases with time. The calibration of autobiographical flashbulb memories is mixed at this juncture (see Neisser 1991; Weaver 1993). However, an interesting study of autobiographical memory suggests the presence of both miscalibration and bias. Bahrick, Hall, and Berger (1996) showed that memory for high school grades is moderately accurate but that distortion is in the direction of higher performance. More important, recall accuracy and confidence are poorly correlated. In fact, confidence is correlated with upward distortion such that more confident answers were associated with higher levels of upward bias.

Although prior study of memory calibration has not been highly sensitive to the consumer context, the implications are evident (e.g., Mazumdar and Monroe 1992). Nostalgic recollection may influence product evaluation; overconfident assessments of one's task performance may inhibit corrective or protective action. In addition, as we elaborate below, confidence in the accuracy of memory for the informational basis of a decision may have an indirect effect on decision confidence.

3. Unconscious, Subliminal, Implicit, and Incidental Learning

Under a variety of names and using a variety of experimental paradigms there is now a large body of evidence that people can learn at least some types of information at very low levels of conscious awareness and control (for reviews of these areas see Bargh and Chartrand [1999]; Berry [1996]; Schacter [1987]; Seger [1994]; Underwood and Bright [1996]). Almost by definition, people should have little or no metaknowledge of unconscious cognitive processes. More specifically, intuition suggests that people should have low confidence for information that they do not know they have. Subjects often believe that they are guessing or report that they are choosing answers because "they feel right" (Berry and Dienes 1993). It is surprising, however, that there is almost no research that directly tests this intuition by measuring confidence as part of an unconscious learning experiment. Exceptions are provided by Dienes et al. (1995) and Chan (partially reported in Berry [1996]), who investigate metacognition in an artificial grammar paradigm, and by Cheeseman and Merikle (1984, 1986), who explore subliminal perception. Across these studies there are two general findings.

First, confidence is frequently uncorrelated with accuracy for implicit learning tasks. Moreover, when the same task is based on explicit learning, a correlation is observed. Second, when people report that they are literally guessing, their performance is well above chance. This outcome is consistent with results from classic calibration experiments and is likely to be due at least in part to regression-to-the-mean effects. Unfortunately, these studies use confidence ratings rather than subjective probabilities; therefore, we do not know whether subjects were also underconfident for nonguessing judgments. However, in sensory discrimination tasks (which include many near-threshold discriminations for which subjects report literal guessing), subjective probability experiments almost always reveal strong underconfidence (Juslin and Olsson 1997). We note that these tasks are relatively difficult (e.g., the percent correct is less than 80 percent), so these results are the opposite of the hard-easy effect typically found for general knowledge questions. Both unconscious learning tasks and sensory discriminations are experienced by subjects as judgments based on feelings that are difficult to articulate. Thus, some authors speculate that lower forms of cognition (such as implicit learning and sensory discrimination) will generally be characterized by underconfidence—in contrast to higher forms of cognition, which tend to exhibit overconfidence. In general, however, research in all domains confirms that when people feel that they are literally guessing, they are likely to be underconfident. From the perspective of consumer decision making, whenever people face choices among near equally attractive alternatives they are likely to be underconfident in their choices.

4. Summary

Are consumers overconfident when they remember the past? Most research supports this claim for both general knowledge and memory for specific events. The sources of this miscalibration are the subject of intense debate (particularly in separating cognitive biases from the effects of statistical error). Nonetheless, these robust empirical effects suggest that consumers are at risk when using memory-based information because sometimes they will think they know more than they do, because their confidence will sometimes be unrelated to accuracy, and because they will sometimes rely on the verbally stated confidence of others. Conversely, there is some basis for optimism insofar as in many situations confidence is correlated with accuracy and over/underconfidence is small. One clear challenge for future research is to identify the situational and personal characteristics that predict the degree of miscalibration. Somewhat ironically, research also suggests that consumers will be underconfident when they are relying on intuition and feel that they are just guessing. Although we certainly do not recommend that consumers rely on guessing more frequently, we do believe that the presumed benefits of laboring over a decision need scrutiny—a conclusion that is reinforced in the next section.

B. Interpreting the Present

In this section, we consider the extent to which currently held beliefs (in contrast with recalled facts and events) are well calibrated. Uncertainty may be reflected in confidence about a particular attribute or option or in the degree of difference among competing options.

1. Belief Polarization

General Beliefs. Although we are inclined to praise decision makers who are thoughtful and introspective, it is becoming increasingly clear that decision quality is not always well served by mental exercise. Initial research conducted largely by Tesser and his associates demonstrated that even brief periods of thought could cause impressions to grow more extreme (see Tesser [1978] for a review). However, the normative implications were not always obvious because a correct degree of attitude extremity is difficult to define and because the effects of thought appeared to be transitory. A possible exception was provided by Langer (1975), who showed that optimistic expectations about a desirable but very low probability chance event increased over the course of a few hours, presumably as a function of thought.

Conceptually similar but more interventionist research has taken a step toward addressing the normative implications of thought. Specifically, this research has examined the effects of instructions to engage in imagination and explanation of a hypothetical outcome prior to judgment (see Koehler [1991] for a review). Consistent with Tesser's results, these studies generally find that thought leads to more extreme beliefs, preferences, and predictions and, in some instances, prompts judgments that are clearly unjustified. Again, however, the connection to miscalibration is indirect, although there are two reasons to suspect that overconfidence was operative. First, as we describe in a subsequent section, the processes associated with polarized thought are processes that are also likely to produce overconfidence. Second, there is both direct and indirect evidence favoring an overconfidence effect. The direct, albeit preliminary, evidence is provided by Sieck and Yates (1997), who report that self-justification (explanation) of a choice prior to the decision results in both higher decision quality and greater confidence in the choice but that the former is apparently outpaced by the latter.

Persuasive indirect evidence of overconfidence resulting from thought can be distilled from measures of perceived disparity among choice options. Consider research suggesting that thought may result in perception of more extreme evaluative differences among choice alternatives (Linville 1982; see also Millar and Tesser 1986). When the dimensions of a set of options are uncorrelated, individuals who think simply (i.e., consider only a few dimensions) perceive greater dispersion across the options than consumers who adopt a more complex thought process. Insofar as complex thought is prompted by expertise as well as motivation, this suggests that novices may perceive greater dispersion than will experts and that all consumers may perceive greater dispersion under low-involvement conditions (but see Lusk and Judd 1988); insofar as complex thought is appropriate, simple thought will result in exaggerated perceptions of dispersion; and insofar as the exaggerated perceptions of dispersion occur along an evaluative dimension, consumers will be overconfident of the superiority of their selection.

Specific Beliefs. More to the point is recent research investigating perceived differences along specific dimensions of competing options. Russo, Medvec, and Meloy (1996; Russo, Meloy, and Medvec 1998) focused on predecisional perceptions and found that perceived differences in attribute attractiveness between two alternatives are exaggerated when a weak (and irrelevant) prior inclination for one of the options exists. Moreover, they found an inertial effect wherein preference for an option on an initial dimension leads to preference for that same option on subsequently inspected dimensions (see also Sniezek, Paese, and Switzer 1990). That is, evaluative interpretation of subsequent information is "distorted" in the direction of the currently favored option and, moreover, the distortion is systematically related to confidence in the superiority of the favored option. A related finding by Holyoak and Simon (1999) shows how desire for a consistent and coherent decision will prompt decision makers to view ambiguous evidence as increasingly supportive of an ultimately favored decision as it is being made. Again, confidence in the correctness of the emerging decision also increases through the decision process, and confidence in the ultimate decision is extreme. Unlike the Russo et al. (1996, 1998) findings, this

effect occurs even when no option is endowed by irrelevant cues. The information provided by Holyoak and Simon (1999) was truly ambiguous, however, as demonstrated by an equal split among subjects regarding the correct option. Given the high confidence exhibited by nearly all subjects, many were clearly miscalibrated.

On the postdecisional side, Svenson and Benthorn (1992) showed that perceived differences between competing alternatives on important dimensions increase over time, but only when the chosen alternative is superior on those dimensions. These results are not inconsistent with the case of unsophisticated thought described previously. Just as failure to think in complex ways can result in exaggerated perceptions of interalternative differences, so too can biased thought. Both should result in less internal conflict over the superiority of the chosen or to-be-chosen alternative—an outcome that may be viewed as a form of overconfidence.

An explicit demonstration of the effects of thought on choice optimality and confidence is provided by Muthukrishnan (1995). In his research, consumers processed information about a target brand, evaluated and/or directly experienced the product to varying degrees, and then chose between the target brand and a dominating alternative. Results showed that higher levels of elaboration and experience led to both a strong preference for the dominated target and higher confidence in the decision.

Whereas the preceding research shows that thought can create overconfidence by exaggerating the differences between options, there is evidence to suggest that thought can even reverse preferences in a counterproductive way. Wilson and his colleagues (e.g., Wilson et al. 1989; Wilson and Schooler 1991) have shown that thought about the reasons for the relative superiority of one option over another can lead to choices that do not correspond to one's true preference. The favored explanation at this juncture is that thought recruits reasons that may not be representative of the spontaneous reasons that otherwise would drive choice, particularly when the decision maker's preferences are not firmly established (Levine, Halberstadt, and Goldstone 1996). Although the shift in preference may be transitory, the consequences may be long lasting. A choice that reflects only momentary preference will later be regretted; thus, novice consumers who think deeply immediately prior to purchase may experience long-term dissatisfaction with their decision (Wilson et al. 1993). Moreover, such instability of preference does not appear to be accompanied by declines in confidence. In fact, people who think about reasons report the same level of confidence in their decisions as do control subjects and, because their accuracy is lower, exhibit relative overconfidence (Wilson and LaFleur 1995; see also Wilson et al. 1989).

2. Belief Validity

Belief overconfidence also arises when consumers possess firmly held beliefs that are inaccurate. There are at least three independent lines of evidence that point to lack of sensitivity to the validity of one's beliefs. First, research on inference making suggests that unjustified inferences exert a powerful influence on judgment. Indirect evidence is provided by Broniarczyk and Alba (1994), who showed that consumer choice may be driven by unsubstantiated inferences even in the presence of other highly predictive cues. In fact, Levin, Chapman, and Johnson (1988) argue that the overconfidence effect may be driven in part by people's uncritical acceptance of their self-generated beliefs. More direct evidence is provided by Sanbonmatsu, Kardes, and Sansone (1991), who showed that, among nonexperts, confidence in one's inferences grows over time. As we will argue later, several illusions of knowledge are attributable to the inability to identify the origin of a belief. Sanbonmatsu et al. (1991) suggest that inferred beliefs increasingly are treated as objective information and therefore are not tempered by uncertainty. Koehler et al. (1996) reported an independent lack of circumspection on the part of decision makers. Their subjects made trait inferences based on valid, related cues but grossly exaggerated the predictiveness of the cues. Consequently, the inferences were too extreme and the confidence intervals surrounding them were too narrow.

A second line of evidence suggests that people are also insufficiently vigilant about the validity of externally generated beliefs. Gilbert (1991) has argued that people are biased to treat incoming information as valid rather than of indeterminate validity. Thus, proper identification and encoding of an assertion as invalid requires additional cognitive activity. When resources are constrained or attention is diverted, correction may not occur and false beliefs may result. Further, retrospective correction may be only partially successful. If additional inferences are spawned from the invalid belief prior to its correction, only the original belief may be rectified (Gilbert and Osborne 1989).

A third and longer line of research further illustrates consumer vulnerability to misperceived truth. It is now well accepted that an increase in the perceived familiarity of an assertion or opinion results in a corresponding increase in its perceived validity (Hasher, Goldstein, and Toppino [1977]; see Brown and Nix [1996] for a recent discussion). Familiarity-based illusions of truth are particularly seductive because the source of an assertion's familiarity can be difficult to identify (Begg, Armour, and Kerr 1985) and because the actual validity of the assertion may be nonobvious (Hawkins and Hoch 1992). Regardless of whether the illusion of truth arises unconsciously or through a lack of corrective action as in Gilbert's (1991) paradigm, overconfident beliefs obtain insofar as increases in perceived validity are equated with decreases in uncertainty. Moreover, insofar as beliefs are precursors to decisions, misperceived validity will result in overly confident decisions.

3. Summary

Evidence regarding the miscalibration of currently held beliefs appears consistent across paradigms. Thought about an object can lead to more extreme but unjustified beliefs, and biased thought can lead to greater perceived dispersion among alternatives and greater perceived differentiation

among attributes. Consequently, people may develop greater but unjustified confidence in their decisions, including their poor decisions. Finally, there appears to be a pervasive lack of sensitivity to the validity of one's beliefs, with the error occurring in the direction of overestimated validity.

C. Predicting the Future

Consumers must not only assess the present state of their own knowledge but also must make forecasts regarding future purchase and usage conditions. Within the classic calibration paradigm this issue has been addressed by substituting forecasting questions for general knowledge questions. A priori, one might anticipate less overconfidence in the case of forecasting because of the uncertainty that accompanies any future event (Ronis and Yates 1987). Unfortunately, the evidence is mostly ambiguous, partly due to methodological concerns. A direct comparison of general knowledge and prediction problems requires equivalence in topic and item difficulty, which may be impossible to achieve. The investigation most sensitive to this problem suggests that prediction can suffer from overconfidence, albeit at a slightly lower level than postdiction (Carlson 1993). We later will discuss process-related reasons to expect differences between general knowledge and prediction. At this point it is sufficient merely to note that the relevance of calibration is limited neither to assessments of current knowledge nor to any particular paradigm. Some of the most important consumer decisions involve future events that are difficult to predict (e.g., interest rates, stock prices); nonetheless, recognition of randomness seems to elude many consumers (see Gilovich 1983; Paese and Sniezek 1991)—sometimes with severe consequences.

1. Personal Outcomes and Ability

Optimism and Desirability. One step in the direction of generalizability is the application of calibration measures to forecasts of personal events rather than knowledge of, or memory for, facts. Consumers often must make forecasts not only about the world (e.g., "This shirt will be marked down by 50 percent in a month") but also about themselves. Such forecasts range from the mundane (e.g., "I'll be out of the store in five minutes") to the momentous ("I can afford this house because my job is secure"). Preliminary evidence suggests that personal forecasts may be overconfident, particularly when the events are difficult to forecast and are viewed as desirable (Pulford and Colman 1996). The desirability bias is consistent with an extensive body of research indicating the existence of unrealistic optimism regarding the probability of personally relevant outcomes (Perloff and Fetzer 1986; Weinstein 1980). More generally, Ganzach and Krantz (1991) report evidence for a leniency effect in which predictions become more positive as uncertainty increases. The implications for consumers are powerful insofar as overconfidence or optimism lowers vigilance toward risks of hazardous consumption, creates inappropriate expectations regarding the likelihood of engaging in successful preventative behavior (e.g., dieting, smoking cessation), or suppresses contingency planning (see Metcalfe 1998; Vallone et al. 1990).

In fact, the planning fallacy represents a specific expression of optimism. Tasks often take longer to complete than the time allotted to them. As Buehler, Griffin, and Ross (1994) demonstrated, individuals are more likely to imagine progress toward a goal than potential impediments. For present purposes, it is noteworthy that this phenomenon exemplifies not only poor prediction but also poor calibration. Buehler et al. (1994) report a sizable discrepancy between the level of on-time completion and confidence that a task would be completed on time. Direct interventions to improve accuracy have produced mixed results (Byram 1997; Taylor et al. 1998). The consequences of the planning fallacy are well known in the context of government spending and public works. The consequences at the individual consumer level involve fewer billions of dollars but can be personally severe, such as when people make overly confident assumptions in the course of devising a financial plan.

Hubris. An interesting aspect of the robustness of optimism is that it appears across a range of events that vary in the degree of control that can be exercised by the individual. On one side of the spectrum, people are optimistic about the outcomes of sporting and economic events over which they presumably have no control (Babad and Katz 1991; Olsen 1997). Calibration also is unimpressive with regard to skilled behavior (see Mabe and West 1982). People overestimate their ability to perform skilled tasks (Harvey 1994) and succeed in competitive situations (Camerer and Lovallo 1999; Cooper, Woo, and Dunkelberg 1988; Larwood and Whittaker 1977). Most relevant to consumer behavior, people are overconfident in their ability to make appropriate choices. Stone (1994) asked subjects to make choices without feedback and estimate the quality of their decisions vis-à-vis others who were making similar decisions. Expectations of success were manipulated. Results showed a high correlation between expectation of success and perceived performance but a low correlation between perceived and actual performance. Overconfidence was highest when expectations of success were highest. In general, people appear to overestimate their desirable abilities and traits vis-à-vis others when the situation provides latitude to do so (Alicke 1985; Dunning, Meyerowitz, and Holzberg 1989; Jones, Jones, and Frisch 1995; Svenson 1981). Moreover, people possess self-serving interpretations of the determinants of success, as reflected in their beliefs about the value of their idiosyncratic traits for achieving particular goals (Kunda 1987).

These results are not inconsistent with evidence gathered from a very different paradigm involving rule learning and usage. In a probabilistic world, perfect prediction is not achievable over the long term. Research on multiple-cue probability learning has shown that forecasters search for a deterministic rule for success, even when informed of its unachievability (Brehmer 1980; Brehmer and Kuylenstierna 1978), and tend to believe that they can outperform a rule

that others have been unable to outperform (Arkes, Dawes, and Christensen 1986). For our purposes the effect of expertise studied by Arkes et al. (1986) is especially enlightening. Experts performed worse because of lower usage of the rule (and presumably greater confidence in their idiosyncratic abilities) yet perceived that they had performed better than did subjects with less knowledge (see also Powell 1991).

If overestimation of one's deep competence arises in situations in which it is impossible to be right without the aid of luck, it should also obtain when it is impossible to be wrong. As we note later, calibration should improve with the receipt of feedback from the environment. However, the environment does not always provide discriminating feedback. For example, in parity product classes in which the level of product performance is high, any choice produces a good outcome. Insofar as assessment of one's decision skills is driven by the quality of decision outcomes (Jones, Yurak, and Frisch 1997), self-confidence should increase irrespective of true ability.

Finally, on a more speculative note, hubris may extend to the general question of one's ability to make rational decisions. Behavioral decision research has been depressingly successful at uncovering a plethora of human decision biases. Some biases are too subtle to be corrected by even the most vigilant decision maker. Other biases achieve a level of awareness that may prompt efforts to take corrective action. Wilson and Brekke (1994) reviewed various aspects of "mental contamination" and provide some preliminary evidence that people overestimate their ability to avoid it. We speculate, based in part on evidence reviewed above, that it would be unsurprising to find overconfidence in one's immunity to the biasing influences in life. Indeed, embedded throughout Nisbett and Wilson's (1977b) critique of verbal reports is informal evidence that people hold confident but erroneous beliefs about the determinants of their own decisions. Consider the specific case of halo effects. Some individuals will strongly deny the influence of a global impression on their evaluations of specific attributes and insist instead that any influence of one on the other flowed in the opposite direction (Nisbett and Wilson 1977a).

Taken together, these findings suggest that miscalibration applies to personal assessments and generally exerts a negative effect on decision making. We offer this conclusion as a stylized fact while readily acknowledging the existence of boundary conditions and moderating factors. Consumers vary in terms of perceived self-efficacy, and not everyone is imbued with a rosy outlook on life. In fact, it appears that depressed individuals may be the best calibrated when the prediction involves desirable events (see Taylor and Brown 1988). Also, in some situations optimism may have beneficial effects that offset the negative effects of inaccuracy. For example, several researchers have suggested that optimism may reduce fear and enhance self-esteem, thereby improving decision making in emotional or anxiety-provoking situations (Metcalfe 1998; see also Luce 1998).

2. Memory, Learning, and Comprehension

A more specific case of personal assessment concerns one's ability to learn, remember, and understand information. Whereas knowledge calibration may be framed in terms of what we know about what we know, memory calibration refers to what we know about what we can remember and comprehension calibration refers to what we know about our ability to understand. Parallel streams of research have grown around each of these aspects of knowledge, much of which involves some form of forecasting.

Metamemory. Consider first the issue of memory. Classic metamemory research falls into two basic paradigms: feeling of knowing (FOK) and judgment of learning (JOL). Feeling-of-knowing judgments involve predictions about one's ability to identify the answer to a memory inquiry given that free recall of the answer has failed. The familiar tip-of-the-tongue phenomenon is a special case. One may not be able to generate the exact answer but nonetheless have a strong feeling of knowing or a confident belief that easy identification will result on a recognition test. For this reason the FOK paradigm is somewhat removed from most consumer memory problems, but it is interesting to note that research on FOK has found that the accuracy of such judgments generally is not particularly impressive (see Costermans, Lories, and Ansay 1992; Perfect and Hollins 1999).

On the other hand, JOL is quite relevant whenever it is important for the consumer to assess the extent to which information learned at time t will be retrievable at time $t + 1$—as when a consumer asks "Will I remember the recipe being described on this show?" or "Will I remember the salesperson's instructions about installation when I get home?" Research on JOL suggests that people are poorly calibrated but that some situations show worse calibration than others. For example, judgments about specific pieces of information (typically words on a word list) are overconfident, whereas judgments about overall memory for an array of information are better calibrated (Mazzoni and Nelson 1995). Also, JOLs are moderately calibrated only when made immediately after information exposure for a later memory test; JOLs made just prior to the memory test are very predictive of memory performance (Keleman and Weaver 1997; Nelson and Dunlosky 1991). Presumably, implicit retrieval attempts made after a delay provide useful cues about ultimate performance; when the JOL must be made upon receipt of the information, however, the availability of information is not diagnostic of later performance (see Benjamin, Bjork, and Schwartz 1998; Dunlosky and Nelson 1997). Unfortunately, many consumer-relevant JOLs, as in the preceding examples, are likely to occur at the time of learning. Moreover, recent research employing stimuli that are less sterile than typically used in JOL studies reports an almost complete lack of correspondence between memory predictions and memory performance (Pritchard and Keenan 1999).

Calibration of Comprehension. Consider next the re-

lated case of comprehension. (In some studies, comprehension and memory are difficult to distinguish because the comprehension questions tap recall rather than integration of concepts.) Proper assessment of one's comprehension is a valuable cue for the need to engage in additional study. If consumers are miscalibrated about their level of comprehension, the result may vary from mild dissatisfaction to egregious mistakes—particularly if miscalibration is in the direction of overconfidence. For example, overestimation of one's ability to install or use computer software or assemble a child's bicycle can lead to exasperation; overestimation of one's comprehension of an insurance or medical policy can lead to catastrophe. Consumer research on the calibration of comprehension is virtually nonexistent, with the possible exception of some suggestive findings in the direction of overconfidence (Chestnut and Jacoby 1980).

Conclusions reached outside the consumer context have evolved over time. Consistent with the evidence provided thus far, initial research suggested a lack of calibration, with a tendency toward overconfidence. Glenberg, Wilkinson, and Epstein (1982) found that subjects maintained a high level of confidence in their self-reported comprehension of a text despite an inability to detect direct contradictions in the text, even when instructed to find contradictions. Glenberg et al. (1982) referred to this behavior as the "illusion of knowing." Subsequent research reported a striking finding, namely, a virtually complete absence of calibration (Glenberg and Epstein 1985, 1987; Glenberg et al. 1987). The typical paradigm in these investigations requires subjects to process stimulus texts and then predict performance on subsequent comprehension questions. The preferred measure of miscalibration essentially represents a correlation between confidence in one's future performance and one's true level of performance. Thus, these initial reports of calibration produced correlations that did not differ significantly from zero. In addition, these studies suggested that people base their predictions of performance not on the text itself but rather on their self-assessed level of expertise in the topic addressed by the text. Subsequent research has moderated these claims to some extent. More valid testing of calibration reports correlations that rise above zero (Weaver 1990), and factors other than mere domain familiarity have been implicated in the confidence judgment (Maki et al. 1990; Maki and Serra 1992; Morris 1990).

Unfortunately, it is not possible to generalize directly to the consumer context. The level of calibration reported in prior research varies significantly as a function of type of processing, the nature and difficulty of the message, the nature of the test questions, performance expectations, and the temporal relationships among the study, prediction, and test phases (see, e.g., Maki 1995, 1998a, 1998b; Weaver and Bryant 1995). Moreover, unlike basic text comprehension, consumer research must conditionalize its results on two pragmatically important factors. First, the existence of a correlation that is statistically above zero but low in absolute size may provide small comfort if the decision facing the consumer is important. Second, although the relationship between a consumer's predicted and true level of comprehension is informative, it is also important to determine the extent to which a bias exists toward over- or underconfidence. The implications for consumer welfare are unlikely to be evenhanded.

Finally, it is important to recognize that comprehension is not restricted to text-based stimuli or, moreover, to one's own level of comprehension. Marketing interactions often (and should) involve assessments of the extent to which a receiver of information (e.g., a consumer) is comprehending the data generated by a provider (e.g., a salesperson). Evidence about such interactions is especially sketchy, but there is reason to believe that both parties may overestimate the true level of comprehension or skill transfer (see Bjork 1994).

3. Assessments of Others

Consumers often consume to benefit others, and consumer researchers have become sensitized to the importance of predicting the preferences of others. Given that consumers can experience difficulty in forecasting their own preferences (Kahneman and Snell 1992; Simonson 1990), it comes as no surprise that prediction of others' preferences is not always accurate. Investigation of the accuracy of forecasts of others' attitudes and behaviors has largely been concerned with the appropriateness of using one's own attitudes and behaviors as a basis for judgment. And, until recently, an extensive body of research had warned against the dangers of projection (Marks and Miller 1987). As embodied in the false-consensus effect, it was believed that inaccurate forecasts result from overestimating the extent to which one's own views generalize to others. It is now becoming apparent that it may be useful to engage in projection, particularly when other grounds for prediction are lacking, even if some overprojection does arise (Dawes and Mulford 1996; Hoch 1987, 1988; Krueger 1998). Of course, projection is least advisable when the forecaster is an outlier.

The question of confidence has not been salient in the study of projection. However, two recent studies show that, whatever the merits of projection regarding accuracy, forecasts of others' preferences may nonetheless be overconfident. Dunning et al. (1990) showed that judgments of others on a variety of dimensions are overconfident, and the level of overconfidence is particularly pronounced when forecasters make predictions that run counter to the base rate. Swann and Gill (1997) extended this finding by showing not only that our assessments of others' preferences, personalities, and histories are overconfident but also that the level of overconfidence is directly related to the length and richness of the relationship between the assessor and the assessee. Swann and Gill (1997) infer that for many judgments, closeness to the target enhances confidence but not the accuracy of the assessor's judgments. This result is interesting in the context of West's (1996) finding that large increases in agent learning can accrue quickly. It seems likely that, as in other demonstrations of overconfidence, the ease with which information can be learned will deter-

mine the level of miscalibration. Moreover, as we have noted, learning (or accuracy) and overconfidence are not mutually exclusive.

4. The Effect of Temporal Perspective

Before concluding this discussion of future events, it is important to note that calibration of forecasts may be sensitive to the imminence of the forecasted event. In some instances of miscalibration the phenomenon is more observable when the to-be-judged event is temporally distant. For example, overconfidence may be higher for more distant events than for imminent events, especially for events that are desirable (Milburn 1978; Wright and Ayton 1992). Similarly, people are more (over)confident of being able to accomplish a task (or to have been able to accomplish a task) the more distant the task is from the time of the assessment (Gilovich, Kerr, and Medvec 1993). Distant forecasts are more ambiguous and may allow greater latitude in creating an optimistic interpretation of a scenario. Accountability for distant forecasts is also less salient, and rigorous consideration of goal-inhibiting constraints or the effort required to attain favorable outcomes may be more easily ignored (Gilovich et al. 1993; Sanna 1999). Of course, proximity to the expected time of an event or the execution of a task may be associated with higher calibration simply because self-awareness of one's ability may be greater, as noted in the cases of meta-memory and comprehension (Mazzoni and Nelson 1995). On the other hand, overconfidence also may increase over time because of underappreciation of one's mounting cognitive limitations (see below).

5. Summary

As in the case of the past and present, predictions of the future also appear vulnerable to miscalibration. Research on personal outcomes and abilities provides the most consistent evidence for overconfidence; predictions about the preferences of others is preliminary but offers intriguing hints of miscalibration. Predictions of learning and comprehension also demonstrate miscalibration but are rife with methodological qualifications and frustratingly subjective assessments of whether the demonstrated level of calibration should be considered encouraging or discouraging.

III. EXPLANATIONS OF MISCALIBRATION

At the beginning of this article we described several sources of miscalibration: cognitive biases, sampling error, response error, and response biases. The latter three are quite general and contribute (at least to some extent) to biased metaknowledge in most situations because random error is ubiquitous and most experimental measures require specific response formats. Cognitive biases, however, are more idiosyncratic and depend on specific models of information processing. In this section we place our focus on cognitive processes. Within the classic subjective probability paradigm, cognitive processes are often suggested but rarely tested rigorously (Keren [1997]; Koehler [1994]; McClelland and Bolger [1994]; see Dougherty et al. [1999] for an exception). Outside the classic paradigm a wide variety of processes have been uncovered that should lead to miscalibration, even though the link is not always explicitly made. We make no attempt to provide an isolated explanation for each phenomenon described, for it likely that many instances of miscalibration are multiply determined. Instead, we describe a variety of processes and indicate their relevance to particular phenomena wherever appropriate.

A. Failures of Memory

The confidence one places in a judgment will be determined in part by the balance of supportive and nonsupportive information relevant to it. In this and the following sections we consider how consumers may develop a false sense of confidence due to biased consideration of evidence. In some instances the decision inputs are biased by virtue of memory processes; in other instances the external world is not searched evenly or completely; and in yet other instances the evidence is tainted by interpretative processes. We refer to evidence in the broad sense. Consumers may fail to consider (1) attribute information when making a brand choice, (2) brands or other problem solving alternatives when deciding on a course of action, (3) predictive cues when making a forecast, (4) motivations and other causal theories when making attributions about a firm's behavior, and (5) impediments to their own successful achievement of a goal.

1. Distorted Memory

Memory is inaccurate, but memory is not randomly inaccurate. Memory often is distorted in a top-down fashion in the direction of predispositions existing at the time of retrieval. The relevant predisposition in many consumer contexts is a prior judgment (see Arkes and Harkness 1980; Ricchiute 1998). Insofar as memory is distorted in the direction of a prior judgment, confidence in the judgment should increase because supportive information will appear more plentiful and consistent than is truly the case. As noted earlier, even misremembered information may be held with high confidence when it is consistent with the retrieval bias (Moeckel and Plumlee 1989). Such miscalibration may occur because the information is constructed on the basis of the retrieval bias and confidence is based on the high degree of plausibility of the information given the retrieval bias (Wagenaar 1988).

The problem posed by distorted memory is that people are unaware of the distortion. Inferences may not be identified as such, and previous beliefs may not be retrievable. Hence, confidence in one's current beliefs cannot be tempered by knowledge of their source.

2. Incomplete Memory

Memory is often incomplete, but it is not randomly incomplete. As in the case of distortion, accurate memory also may be guided by prior judgment. Confidence in a previous decision will be overstated if memory for decision-relevant information is skewed in the direction of decision-consistent facts. Dellarosa and Bourne (1984) provide evidence for such a bias, which they explain in terms of imbalanced cuing of decision-consistent facts by the decision itself. This explanation, however, presumes either a motivational or logical bias to retrieve supportive facts or, as Dellarosa and Bourne (1984) prefer, a retroactive reorganization of facts in memory such that supportive facts are linked to the decision after the decision has been made. As we describe below, motivational and logical biases may indeed result in miscalibration. In terms of memory, a simpler explanation is suggested by Hoch (1984), who also reported biased retrieval of decision-relevant information but attributed it to order effects rather than higher-order cuing. If decision makers first recall supportive information, recall of unsupportive information may be reduced as a result of output interference. An attractive feature of Hoch's explanation is generalizability. That is, output interference is not reliant on a previous decision. It applies equally well, as demonstrated in Hoch's experiments, to miscalibrated forecasts of future behaviors and outcomes and to decisions not yet made.

Irrespective of the precise explanation of Dellarosa and Bourne's (1984) findings, there is ample evidence to support the operation of both outcome or decision cuing and interference. Consider the phenomenon of hindsight bias. The failure to be surprised by outcomes—or the feeling that one "knew it all along"—could result in a feeling of omniscience. Thus, one does not become better calibrated over time because the feedback essential for good calibration is not chastening. Although a definitive account of hindsight bias has been elusive, Hawkins and Hastie (1990) argue that the most viable explanation involves rejudgment, wherein recall, reconstruction, and evaluation of evidence for the known outcome is guided by the outcome itself. Thus, retrieved knowledge pertinent to the question is biased in favor of the outcome (see also Erdfelder and Buchner 1998; but see Winman, Juslin, and Bjorkman 1998).

A larger and more diverse body of evidence supports the notion that decisions may be miscalibrated because retrieval of relevant information is thwarted by the interfering effects of other salient information. The clearest examples beyond Hoch's (1984) demonstration can be found in the consideration of decision alternatives. In this case, miscalibration stems not from biased assessment of a particular course of action but in misestimating the alternative courses of action. Overconfidence results from the belief that an option has been selected from the complete set of possibilities when in fact superior alternatives have not been considered. For example, consumer research has shown that failure to generate a complete set of decision alternatives results in choices that would have been made differently had a complete set of options been present (Alba and Chattopadhyay 1986; Nedungadi 1990; Posavac, Sanbonmatsu, and Fazio 1997).

We are left to speculate about these consumers' level of calibration, inasmuch as no measure was taken that even approximated decision confidence. However, other research is suggestive. First, results obtained from the fault-tree paradigm demonstrate that people are captured by the "options" that are explicit and therefore underestimate the number and/or probability of occurrence of unspecified options (e.g., Fischhoff, Slovic, and Lichtenstein 1978; Russo and Kolzow 1994). An exaggerated effect occurs when people generate a set of options and then must estimate the likelihood that the set of options they have generated includes the optimal alternative or that the set encompasses the complete set (Gettys, Mehle, and Fisher 1986; Gettys et al. 1987; Mehle 1982; Mehle et al. 1981). Overconfidence obtains when people overestimate the probability that the best option is contained in the set or overstate the likelihood that a particular option is correct. A formalized treatment of a special case of this phenomenon can be found in Support Theory (Tversky and Koehler 1994), which argues that the subjective likelihood of a particular hypothesis will be influenced by the extent to which decision makers consider or unpack all aspects of the hypothesis.

The most persuasive evidence for this cause of miscalibration is provided by studies that directly measure confidence and calibration. Consistent with the preceding logic, Kochler's (1994) results suggest that overconfidence subsides as decision makers consider alternative hypotheses (see also Hoch 1985; Kuhn, Weinstock, and Flaton 1994; Peterson and Pitz 1988). Analogous findings are common in other decision contexts. For example, a particular decision may be biased by the attention-capturing effects of salient decision variables (Glazer, Steckel, and Winer 1992; Sanbonmatsu, Akimoto, and Biggs 1993). Variables not presented may remain unconsidered. Insofar as individuals assume they have conducted an exhaustive analysis, overconfidence may ensue. In fact, Sanbonmatsu, Kardes, and Herr (1992) demonstrated that consumers who were less likely to consider important product dimensions were more confident of their decisions. A similar effect can be observed in competitive situations when even trained and motivated decision makers fail to consider the implications of their actions on the contingent actions of their competitors (Zajac and Bazerman 1991). Conversely, failure to reason "consequentially" may lead to underconfidence if consumers fail to realize that the wisdom of a decision will be unaffected by the outcomes of other events (Shafir and Tversky 1992).

The failure to give appropriate consideration to all possible outcome states may underlie inaccurate subjective probabilities of outcomes and overconfidence in causal explanation. The literature does not speak with a single voice, but it is apparent that imagining and/or explaining an outcome increases its subjective likelihood (Anderson and Wright 1988; Carroll 1978; Hirt and Sherman 1985; Levi and Pryor 1987). Specific accounts vary but typically include some notion of availability (Tversky and Kahneman

1973). Perceived likelihood of an outcome varies with the ease with which it can be imagined and the frequency with which it is imagined—irrespective of one's attitude toward it (Anderson 1983b; Anderson and Godfrey 1987; Sherman et al. 1985).

These effects can be pernicious for at least two reasons. First, decision makers appropriately seek causal explanations and are guided in their diagnoses and forecasts by the plausibility and coherence of those explanations (see Einhorn and Hogarth 1986; Jungermann and Thuring 1987; Kahneman and Tversky 1982; Pennington and Hastie 1988; Tversky and Kahneman 1983). However, neither imaginability, plausibility, or coherence of an explanation is perfectly predictive of its validity, yet each may inhibit consideration of alternative theories that may be equally valid (cf. Anderson and Sechler 1986; Gregory, Cialdini, and Carpenter 1982). Second, incorrect explanations may persist even after the data on which they were based have been discredited, either because the theory exists independently of the data (Anderson, Lepper, and Ross 1980) or the discredited data offer the only basis on which a theory can be constructed (Johnson and Seifert 1994). In either case, the explanation is held with greater strength than the data justify (see Anderson 1983a).

B. Attentional Failures and Misweighting of Evidence

As suggested, miscalibration may arise when decision inputs are not optimally incorporated into the decision process. Such errors may occur even when memory constraints have been lifted (see, e.g., Sanbonmatsu et al. 1998).

1. Base Rates

Differential salience and availability of information has been implicated in nonregressive thinking. Overreliance on the extremity of a cue rather that its validity, as reflected in underweighting of base rate information, has been documented across a variety of contexts (see Griffin and Tversky 1992). In turn, several manifestations of miscalibration have been attributed to nonregressive thinking (cf. Kahneman and Lovallo 1993). For example, the planning fallacy has been explained in terms of overweighting of the characteristics of target tasks at the expense of ample past experience with the failure to meet deadlines (Buehler et al. 1994). Overly confident inferred beliefs have been attributed to an overestimation of the strength of relationship between variables, that is, to a failure to temper the extremity of a cue with its statistical predictiveness (Koehler et al. 1996). Confident predictions have been shown to be based on the extremity of a single explicit cue, even when the existence of other valid and uncorrelated cues are easy to imagine (Peterson and Pitz 1986). In interpersonal prediction, overly confident judgment may be based on the richness of nondiagnostic knowledge of the target (Swann and Gill 1997), and overconfident predictions of self and others have been shown to be largely confined to instances in which the predictions run counter to the base rate (Vallone et al. 1990). In addition, overconfidence on difficult questions may be attributable to overreliance on available but biased cues. As Schneider (1995) notes, confidence is not a good predictor of performance on difficult questions perhaps because people are insensitive to their lack of specific knowledge and overly reliant on salient but nondiagnostic cues such as familiarity.

2. Incomplete Generation and Assessment of Evidence

Salience and availability are not the only barriers to calibrated judgment. Miscalibration may obtain even when all relevant information is available or obtainable. In particular, we refer to the unevenness with which confirming and disconfirming evidence is sought, encoded, and processed (see Klayman [1995] for a discussion). Indeed, Einhorn and Hogarth's (1978) theoretical treatment of overconfidence is built largely on people's failure to process hypothesis-disconfirming evidence. As in other aspects of overconfidence, empirical evidence varies in its explicitness. Consider Shaklee and Fischhoff's (1982) demonstration that decision makers will attempt to confirm a leading hypothesis and then truncate the search process, even when alternative hypotheses have been suggested. Although truncated search may be viewed as a measure of confidence, Shaklee and Fischhoff's task lacked a true criterion. However, Wason's (1960) seminal—albeit narrow—examination of confirmation bias may be viewed as an implicit demonstration of overconfidence inasmuch as subjects were asked to cease testing once they were certain that they had uncovered the appropriate hypothesis in a rule-learning task. Numerous subjects prematurely declared awareness of the rule.

Fortunately, demonstrations of insufficient hypothesis testing have been produced within the classic overconfidence paradigm. When presented with general knowledge questions, overconfidence is reduced when decision makers are asked to generate reasons why their response may be incorrect, presumably because such counterfactual thinking does not occur spontaneously (Koriat, Lichtenstein, and Fischhoff 1980; see also Hirt and Markman 1995; Hoch 1985). Similarly, overconfidence is exaggerated when decision makers are asked to assess the correctness of a prespecified response relative to when they are free to consider both response alternatives (Ronis and Yates 1987; Sniezek et al. 1990).

Supportive results also are obtained outside the traditional paradigm. For example, Brenner, Koehler, and Tversky (1996) presented subjects with one-sided or two-sided information in a jury prediction task. Subjects provided with one-sided evidence naturally exhibited less accuracy but also showed greater confidence in their predictions—despite explicit reminders that they were basing their judgments on asymmetric information. Consistent with Koriat et al. (1980), only instructions to consider the likely strength of the unpresented side were effective. Such biasing effects may be sufficiently strong to influence monetary decisions.

Gibson, Sanbonmatsu, and Posavac (1997) showed that asymmetric evaluation of competing sports teams not only biases judgment of an arbitrarily chosen focal alternative but also increases willingness to wager on its likelihood of success (see also Anderson and Sechler 1986). These results are consistent with the direction-of-comparison effect in consumer choice wherein decision makers prefer a focal alternative because they retrieve more favorable information about it than about the nonfocal alternative (Dhar and Simonson 1992).

These results are also consistent with the apparent tendency to reason pseudodiagnostically. It has been argued that people are capable of considering only one hypothesis at a time (Mynatt, Doherty, and Dragan 1993; see also Van Wallendael and Hastie 1990). Hence, a cue may be taken as support for a focal hypothesis even though it is equally true of a competing hypothesis. Although controversy has arisen regarding the extent to which decision makers consider the diagnosticity of information in their judgments (e.g., Doherty et al. 1979; Skov and Sherman 1986), it appears that people are sensitive to the diagnosticity of cues only when the diagnosticity of those cues is easy to determine; when diagnosticity is not obvious, spontaneous consideration of diagnosticity is uncommon even when the opportunity to obtain information about diagnosticity is salient (Beyth-Marom and Fischhoff 1983; Doherty et al. 1996). Consider the case in which consumers receive information but must search and test additional cases in order to assess diagnosticity. Hoch and Deighton (1989) argue that consumers who purchase a brand because of a particular advertised feature may not consider the fact that other brands may have the same feature. Insofar as the feature is perceived to be attractive and unique, confidence in the superiority of the brand may result. Similarly, if a consumer repeatedly samples the same individual brand or the same set of high-priced brands because they provide a positive outcome, beliefs about the superiority of these brands may arise because alternatives are never sampled (see Einhorn and Hogarth 1978; John, Scott, and Bettman 1986).

C. Inappropriate Decision Inputs

Miscalibration may arise when cues are misinterpreted and inappropriately influence decisions.

1. Inferences Based on Task Characteristics

Information Intensity. Confidence that one has or will perform well on a task may be driven by metacognitive beliefs about oneself as well as task characteristics that should affect performance. For example, greater amounts of data should result in more fully informed decisions. When changes in the quantity of data are salient, confidence may vary with information density—even when the data are nondiagnostic or do not enable improvements in accuracy that are commensurate with increases in confidence (see Oskamp 1965). Peterson and Pitz (1988) provide a compelling demonstration by ruling out the possibility that greater amounts of information merely provide people with greater latitude to interpret information in a biased manner. Confidence was more closely associated with level of information intensity when the latter was manipulated within subject than when it was manipulated between subjects. Thus, as people viewed more information, they became more confident. The strength of this effect is illustrated by the fact that the information itself could be interpreted in multiple ways, as indicated by the higher number of hypotheses prompted by higher levels of information.

Task Involvement. Effort and practice are additional cues often associated with performance. As in the case of information quantity, research suggests that both factors produce a dissociation between confidence and performance, again because of elevations in confidence that are unaccompanied by corresponding increases in accuracy (Paese and Sniezek 1991; see also Sieck and Yates 1997). Similar themes were echoed by Fischhoff and Slovic (1980), who argued that even minimal familiarity with a task or tentative rules of performance in the task can produce surprising amounts of conviction, and by Heath and Tversky (1991), who noted that familiarity and knowledge are accompanied by feelings of competence. It would be misleading to suggest that knowledge and competence are unrelated. However, there are circumstances in which tasks are inordinately difficult and expertise is not easily acquired. When cues normally associated with the accumulation of expertise are used indiscriminately, overconfidence may result. Consumers face no shortage of such circumstances. Brand superiority in near-commodity product classes and subtleties in the attributes that discriminate one product from another—as in the case of wine—may be difficult to discern despite high levels of experience, practice, and effort.

These metacognitive beliefs bear a strong similarity to the processes thought to underlie the illusion of control (Langer 1975). Indeed, the illusion of control may be viewed as both a form of overconfidence as well as an explanation of it. Langer argues that perceptions of control will arise in random environments if the task has features normally associated with control (e.g., practice). Insofar as people believe that they exert more control over outcomes than is possible, overconfidence should characterize task performance. Thus, it is not surprising that subsequent research on miscalibration has shown that illusory control may exacerbate overconfidence in the traditional paradigm (Henry 1994) or that overconfident forecasts may be produced by activities that emulate analytic thinking but that can exert no possible salutary influence on actual performance (Kottemann, Davis, and Remus 1994; cf. Gollwitzer and Kinney 1989).

Although prior research has been extreme in its manipulation of control (usually allowing subjects no control) in order to demonstrate the illusion, real-world control over many outcomes exists in varying degrees. Consumer decisions regarding risk (e.g., how much insurance to purchase) pertain to outcomes that offer partial control. Thus, it is also not surprising that miscalibration regarding susceptibility to

risk is related to perceived control over the risk (see Harris 1996; Quadrel, Fischhoff, and Davis 1993). The consumer's error is not in assuming control over the uncontrollable but in overestimating control (or underestimating randomness), which in turn may lead to unrealistic optimism and underreliance on imperfectly predictive decision rules.

Task Domain. Finally, a variant of the control metabelief is the knowledge metabelief. That is, consumers may feel confident about their own domains of expertise. Overconfidence obtains when expertise is delusional or unhelpful (Kleinmuntz 1990; Trafimow and Sniezek 1994). Thus, comprehension and feelings of knowing may be miscalibrated because people base their predictions of future performance on general knowledge of a topic rather than specific understanding of a stimulus (Costermans et al. 1992; Glenberg and Epstein 1987), and underutilization of predictive rules by experts may stem from misestimation of the usefulness of expertise in the context of unpredictability (Arkes et al. 1986). On the other hand, knowledge of one's general ability may improve calibration if the to-be-judged event benefits from use of metaknowledge. For example, there is some evidence that aggregate judgments of performance are better calibrated than individual judgments (Gigerenzer et al. 1991; see Treadwell and Nelson [1996] for a discussion). It has been suggested that the former are based not on the sum of the individual confidences (which may be biased for reasons described above) but rather on a general sense of overall knowledge of the domain, which tends to be low in most calibration experiments (Sniezek and Buckley 1991).

2. Misattribution Effects

Miscalibration has been discussed in terms of misestimation of one's knowledge and misperception of the validity of assertions. Misattribution of information and beliefs has been implicated in each. For example, an early explanation of the classic miscalibration effect offered by Fischhoff, Slovic, and Lichtenstein (1977) involves confusion about the status of retrieved information. Specifically, answers that are constructed from fragments of retrieved information are interpreted as having been recalled from prior knowledge, that is, inference is interpreted as recall. Insofar as people are more confident of their knowledge than their inferences, overconfidence obtains. Carlson (1993) relied on this speculation to justify different levels of calibration between general knowledge tasks and forecasting tasks. If forecasts are better calibrated, it may be due to the reduced level of misattribution; that is, because forecasted events have not yet transpired, inference cannot be interpreted as recall.

This explanation implies that as more operations are needed to arrive at an answer, confidence in the answer should decline (see Allwood and Montgomery 1987). In fact, recent research has shown how misperception of one's knowledge can be influenced by the familiarity of a response—with familiarity judgments requiring little cognitive effort. For example, information made recently familiar may be attributed to long-term knowledge (Begg et al. 1996). More direct evidence pertaining to miscalibration has been produced through the use of general knowledge questions. Although answers to such questions should be based on the amount of information that can be gathered in their support, it now appears that the ease with which the response can be retrieved also plays a large role. Preexposure to a plausible but incorrect answer leads both to fluent retrieval of the incorrect response and confidence in its accuracy (Kelley and Lindsay 1993; see also Chandler 1994). Fluency apparently is interpreted as knowledge, which, in this particular case, leads to confident incorrect responses. In a sense, this outcome may be viewed as a metabias inasmuch as retrieval fluency can be a barometer of knowledge. However, it is an imperfect indicator that will lead to overconfidence when fluency is misattributed to internal knowledge rather than to external sources of familiarity.

The misattribution of familiarity may lead to an illusory sense of knowledge in prediction tasks as well. Feeling-of-knowing tasks require people to forecast their ability to recognize unrecallable answers. Although FOK responses should be based on the partial knowledge that is retrievable (Koriat 1993), overly confident predictions may be based on the mere familiarity of components of the question (Reder and Ritter 1992; Schwartz and Metcalfe 1992).

Judgments will also be miscalibrated if the decision maker is incorrect about the validity of the inputs into the decision. Most research dealing with this issue has focused on the case in which validity is overestimated. Several independent causes of truth overestimation have been probed, all of which involve source misattribution. The classic illustration is the sleeper effect, in which the perceived validity of an assertion increases over time because of the lessening influence of its source cue (see Pratkanis et al. 1988). In the end, a questionable assertion may be viewed as valid because the reason to question its validity is not retrieved and because there is a prevailing bias to treat unquestioned information as valid.

The sleeper effect represents a case of source misidentification. A similar confusion occurs in the memory-misinformation paradigm when eyewitnesses confuse a suggested detail with the true fact—a problem that worsens the more often the misinformation is rehearsed or suggested (Shaw 1996; Weingardt, Loftus, and Lindsay 1995; Zaragoza and Mitchell 1996). Direct research on misperception of validity—in which the perceived truth of an assertion increases with exposure—has implicated source confusion but also relies on notions of familiarity (Begg, Anas, and Farinacci 1992). The latter is reflected in the observation that the bias is driven by increases in familiarity rather than exposure frequency, per se (see Bacon 1979; Hawkins and Hoch 1992). On the other hand, the truth bias is exacerbated by failure to recall the source of the assertion, as in the sleeper effect (Arkes, Hackett, and Boehm 1989; Law and Hawkins 1997; see also Brown and Nix 1996). In this sense, both biases are akin to the "false fame" effect in that fa-

miliarity is misattributed to a factor other than mere exposure (Jacoby et al. 1989).

Outside the domain of the truth bias, much of the discussion of source misattribution has centered on the confusion between external and internal sources (see Johnson, Hashtroudi, and Lindsay 1993). Such confusion may also lead to misperceptions of validity via "imaginal confirmation" of beliefs (Slusher and Anderson 1987). That is, imagined validation of beliefs is confused with external validation, and therefore the total perceived support is greater than the actual level of support. Similarly, high but erroneous confidence in false memories may result from external-internal confusion, particularly when the internal representation is retrieved fluently (Holmes et al. 1998; Jacoby, Kelley, and Dywan 1989; Roediger and McDermott 1995).

3. Inferences about Memory

Memory is a partial reflection of our knowledge, and overconfidence in the state of our knowledge may be influenced by metacognitive beliefs about memory. Gentner and Collins (1981) illustrate a type of memory conceit by showing that an inability to recall information about an assertion predisposes one to interpret the assertion as false. A symmetrical bias may exist wherein a recalled assertion is interpreted as true. Such a bias would be consistent with the sleeper effect and the truth bias. Familiar assertions are more likely to be treated as true; and, by extension of Gilbert's (1991) logic, a recalled assertion should be treated as valid until actively negated.

The malleability of memory may also enable higher-order beliefs to alter current and retrospective perceptions of ability. Substantial evidence suggests that memory for objective information is biased (reconstructed) in the direction of one's beliefs and prejudices (see below). Recent evidence further suggests that memory for prior personal states (feelings, opinions) is biased in the direction of one's current state (Levine 1997; Zwick, Pieters, and Baumgartner 1995). Ross (1989) has argued that, with the current state as a reference point, the prior state is reconstructed based on implicit theories about the direction change should take. (See Kahneman and Snell [1992] and Simonson [1990] for prospective analogues.) With regard to confidence, consider a specific example by Conway and Ross (1984). Subjects trained to improve their study skills recalled their pretraining skills as less developed than they had rated them prior to training; thus, subjects exaggerated the degree of skill enhancement. These results suggest that memory could serve to exaggerate the confidence-distorting effects of practice, effort, and familiarity described above. Cues that lead to unjustified confidence may also lead to distorted impressions of one's previous level of competence. Of course, different theories about these cues or other cues could result in overestimation of one's previous ability and underconfidence.

D. Motivated Reasoning

The discussion thus far accurately reflects the dominant cognitive orientation of research on calibration. However, motivational factors may also play a role. Pertinent to forecasting, for example, the desirability of an outcome has been shown to be negatively correlated with calibration, which may partially account for optimism (cf. Buehler, Griffin, and MacDonald 1997; Wright and Ayton 1989). In addition, there appears to be a strong inclination toward self-enhancement, which may partially explain overestimation of skilled performance (see Sedikides and Strube 1997).

Motivational effects are most evident in the context of information acquisition and belief persistence. As noted, overconfidence may arise if consumers incorrectly believe that they have adequately searched the information environment. Specific motivational forces may vary across consumers and situations. Thus, consumers may attempt to justify small samples in order to avoid search effort (Santioso and Kunda 1991), selectively search information to minimize feelings of dissonance (Frey 1986), engage in confirmatory testing to avoid negative outcomes (Tschirgi 1980), exaggerate a focal cause when being held accountable for a judgment (Sanbonmatsu et al. 1993; see also Lerner and Tetlock 1999), or terminate search prematurely in order to satisfy needs for closure or structure (Kruglanski and Webster 1996; Mayesless and Kruglanski 1987).

Similarly, beliefs may persist in the face of disconfirming evidence as a result of motivation to maintain those beliefs (cf. Kuhn and Sniezek 1996). A relatively subtle avenue to perseverance involves asymmetric consideration of information such that supportive evidence is accepted uncritically whereas threatening evidence is inspected carefully or held to higher standards of proof (Ditto and Lopez 1992; Ditto et al. 1998). In other instances, motivation to maintain one's beliefs may lead to very selective interpretation and active refutation of disconfirming evidence (Edwards and Smith 1996; see also Koehler 1993; Kunda and Oleson 1997; Svenson 1992). When amplified by self-enhancement motives, such behavior may account for unrealistic optimism and irrational persistence in a difficult or impossible task. As Gilovich (1983; Gilovich and Douglas 1986) has demonstrated, people may maintain confidence in their (in)ability despite regular failure feedback through creative interpretation of the causes of their failure, even in tasks that produce random outcomes.

Nonetheless, motivational accounts of overconfidence do not necessarily substitute for cognitive accounts (see Metcalfe 1998). As Kunda (1990) notes, motivational factors may create initial top-down biases that are then sustained by the cognitive mechanisms described throughout this article.

IV. FUTURE RESEARCH

Prior research on judgment and decision making has emphasized the accuracy of people's decisions. An independent line of research has examined the second-order question

concerning the confidence with which beliefs are held. We have tried to illustrate several distinct domains important to consumer behavior in which accuracy and confidence have been joined to provide evidence of calibration. In some instances the evidence is direct; in other instances, calibration is more speculative. In all instances the implications for consumer behavior are significant but the external validity of the paradigms is unproven. Taken together, it is evident that additional research is necessary before strong conclusions can be drawn about calibration of one's own and others' knowledge, forecasts, memory, comprehension, plans, or any other aspect of behavior for which confidence is relevant. Moreover, as we have noted repeatedly, the extensive literature on calibration and calibration-related processes often fails to make direct contact with the world of consumer behavior because the task and/or the stimuli are far removed from everyday experience. This state of affairs is not unusual, and consumer researchers have become adept at adapting paradigms to advance our understanding. Thus, one recommendation is to examine many of the issues discussed above in a more externally valid setting. However, we can easily go beyond this obligatory statement by noting some of the many opportunities to expand the paradigm by building on incipient research and exploring new domains.

A. New Applications

A simple but important task is to pose the calibration question in any arena in which performance can be measured. For example, consider the case of taste or preference forecasting. It appears that the accuracy of our predictions regarding how our preferences will change over time is poor (Kahneman and Snell 1992), yet we know little about confidence in those predictions. Calibration can also be assessed with regard to choice processes. As we have noted, decision makers often lack insight into their judgment processes (Christenfeld 1995; Nisbett and Wilson 1977b). It is unclear, however, whether high confidence is regularly placed in erroneous rationales. Even when particular heuristics are knowingly used, the appropriateness of consumer confidence in the ability of the heuristic to produce a favorable outcome at a predicted level of processing cost is an unexplored issue despite the emphasis placed on strategy selection by consumer researchers (e.g., Payne, Bettman, and Johnson 1988).

B. Confidence and Performance

Optimism and inflated feelings of self-efficacy may have a motivational effect that ultimately results in higher levels of performance (Sherman et al. 1981; Taylor and Brown 1988; Wood, Bandura, and Bailey 1990). It is unclear how these motivational effects influence consumer behavior or how consumers deal with instances in which the outcomes are disappointing (Armor and Taylor 1998).

C. Overconfidence versus Risk Aversion

It is striking that the research on overconfidence rarely refers to the conflicting tendency toward risk aversion. Kahneman and Lovallo (1993) remind us of this paradox, noting the possible neutralizing effect that each may have on the other but lamenting the occurrence of rational decisions that are based on two opposing irrational processes. Given the lack of research devoted to consumer calibration, it follows that the question of how risk aversion interacts with miscalibration requires attention. Perhaps we have been too quick to accept risk aversion as a main effect. A true test of its robustness requires investigation in contexts in which the processes underlying optimism and overconfidence are operative. Similarly, decision confidence should be examined in situations involving real risk.

D. Interpersonal Interaction

Decision making is not always an isolated process. As noted at the outset, there are situations in which the calibration level of another person is of more concern than one's own. Aside from simply relying on the opinion of another, there are many occasions on which a decision is made by the consumer but is influenced by interaction with others. Much of the research that pertains to interpersonal effects understandably has been conducted in group settings with an eye toward organizational decision making. Sniezek (1992) provides a review, from which she concludes that group decisions are also susceptible to overconfidence although perhaps to a lesser degree than individual decisions.

Surely there are consumer contexts that emulate a true group (e.g., egalitarian family decisions), but it seems that there are many other situations that involve interpersonal interaction on a less structured basis, such as in the case of word-of-mouth communication and professional advice. Consumer research has been remiss regarding decisions beyond the individual level, and calibration research in general has not spoken to scenarios that map onto interpersonal consumer decision making. An intriguing exception is provided by Heath and Gonzalez (1995). They examined the ecologically common case in which an individual consults another prior to making an individual decision. Although there is reason to expect overconfidence to be tempered by consultation with another, Heath and Gonzalez (1995) found that such interaction increased confidence without affecting accuracy. Their preliminary explanation is that interaction does not ensure a fair and full give-and-take. If interaction instead prompts one to explain or defend one's own point of view, overconfidence may obtain. We are reluctant to speak to the generalizability of this outcome because the structure and content of interpersonal interaction is likely to vary widely across consumers and situations. In fact, other evidence suggests that overconfidence may be reduced in those instances in which partners exchange differing views in a balanced manner (Allwood and Granhag 1996) and on occasions in which decision makers anticipate future social interaction that may encompass justification of a judgment

(Tetlock and Kim 1987). At a minimum, Heath and Gonzalez (1995) point to the richness of the issue and the need for additional investigation.

E. Evaluation of Others

Consumer decisions frequently are based on the recommendations of others. An obvious question for consumer research concerns the accuracy of the recommendation. An equally obvious question from our perspective concerns the second-order issue of the confidence one places in the accuracy of others. Unfortunately, research is sparse and fragmented. The most recent studies have adapted the classic calibration paradigm to fit the situation in which actors make judgments and observers assess their confidence in the actors' responses. Although results vary as a function of task, it appears that observers show a general pattern of overconfidence that is not unlike that reported in the classic paradigm (Harvey, Koehler, and Ayton 1997; Koehler and Harvey 1997). Evidence from other paradigms suggests that both the perceived ability and credibility of others may be based on cues that are not strictly valid (Bell and Loftus 1989; Jameson et al. 1993). In fact, research on deception suggests that our ability to detect mendacity may be poor indeed (Hyman 1989; Zuckerman, DePaulo, and Rosenthal 1981). Insofar as consumers overestimate their ability to judge another's character, they will be miscalibrated with regard to the validity of information provided by salespeople and therefore vulnerable to interpersonal persuasion attempts.

In some instances, consumer confidence may rise or fall with observations of the behavior of others (see Luus and Wells 1994). Although consensus cues may provide useful input to a decision, the consensus may be wrong and herd behavior may result in undesirable outcomes on a large scale (see Bikhchandani, Hirshleifer, and Welch 1998). Given the importance of the issue of diffusion, it is an appropriate time to examine the effects of consensus on confidence and accuracy.

F. Cross-Cultural Effects

Yates and his colleagues have begun to explore the robustness of miscalibration as a function of nationality (Yates, Lee, and Bush 1997; Yates, Lee, and Shinotsuka 1996; Yates et al. 1989). Preliminary evidence suggests that, at least within the classic paradigm, calibration and accuracy vary between Eastern and Western cultures, with the unanticipated outcome that overconfidence may be greater in Chinese than in American samples. In addition, there is some evidence to suggest that self-enhancing biases that may produce unrealistic optimism vary across cultures in ways that are consistent with the familiar East-West difference in individualism/collectivism (Heine and Lehman 1995). Clearly, however, research on this topic is in its infancy.

G. Lability of Values

Finally, research within and beyond the consumer context has questioned whether people possess deep values on issues of the day (Feldman and Lynch 1988; Fischhoff 1991). The issue has obvious implications for the reliability of consumer surveys (see Fitzsimons and Morwitz 1996; Simmons, Bickart, and Lynch 1993). Insofar as values are not firmly held, it is important to investigate the confidence people have in their opinions and the effect of survey intervention on that confidence. As Fischhoff, Slovic, and Lichtenstein (1980) note, people may be unaware of the lack of firmness of their own opinions and therefore content to articulate top-of-mind responses. However, consistent with our discussion of polarized judgments, people may become committed to these opinions based on little more than the need to maintain consistency and avoid uncertainty (cf. Svenson 1992).

H. Individual Differences

Across Domains. The question of whether people are consistent in their degree of miscalibration has been raised only recently. Nonetheless, there are two preliminary lines of evidence that suggest that miscalibration may have a traitlike quality. First, miscalibration varies across individuals but may be correlated within individuals across very different tasks. Individuals who exhibit overconfidence on a test of general knowledge tend to be overconfident on tests of episodic memory and motor performance (Bornstein and Zickafoose 1999; Klayman et al. 1999; West and Stanovich 1997). Second, miscalibration may be correlated with other measurable traits. For example, the decline in unrealistic optimism observed as a task becomes temporally nearer is not observed uniformly across individuals. Rather, the greatest reduction occurs among those who suffer from low self-esteem (Shepperd, Ouellette, and Fernandez 1996). Given the wide variance not only in the manifestations of miscalibration but also in the underlying processes, truly robust individual differences may be difficult to demonstrate. However, insofar as miscalibration is driven by common processes (e.g., overestimation of personal control), consistent individual difference may emerge. The question merits attention given the pragmatic significance of segmentation in consumer markets.

Expertise. It is fitting to conclude with one of the most important and frequently investigated individual differences in consumer research. The effects of expertise have both theoretical and practical significance; however, research has yet to draw a firm conclusion. For example, meteorologists, accountants, bridge players, and some gamblers appear to be very well calibrated, whereas doctors are not (see Keren 1991; O'Connor 1989; Yates and Curley 1985). Financial analysts appear susceptible to wishful thinking (Olsen 1997), and econometric forecasters are well-calibrated in the very short term but not beyond (Braun and Yaniv 1992). Expert property appraisers exhibit higher overconfidence than do novices because their superior accuracy is not sufficient to

compensate for their even greater confidence (Spence 1996; see also Paese and Feuer 1991). The results are also mixed when subjective expertise is assessed. Higher self-assessed expertise in snooker is associated with better calibration (Wright et al. 1994). However, Bradley (1981) reports that for difficult tasks, self-rated experts are less willing to admit uncertainty and more likely to express confidence in incorrect responses. Additional research may be required to explain these disparate results. One tentative conclusion is that calibration, like expertise, requires diagnostic experience within a domain. Calibration requires experience in the form of specific feedback about the accuracy of one's judgments. It also requires attention to that feedback (cf. Arkes et al. 1987; Keren 1991; Mahajan 1992; Sharp, Cutler, and Penrod 1988). The extent to which consumers receive frequent and unambiguous feedback and then map that feedback onto their judgments in a unbiased manner is an open question. However, there is little cause for optimism. Recent evidence suggests not only that novices perform poorly on many tasks in comparison to experts but are also particularly insensitive to their poor level of performance—perhaps as a result of a dearth of negative feedback, a failure to make appropriate attributions when feedback is provided, and an inability to discriminate good from poor performance in the environment at large (Kruger and Dunning 1999).

V. CODA

A sizable body of psychological research argues forcefully that knowledge exerts a significant and fundamental influence on consumer decision making. What (and how) consumers decide varies as a function of what they know. This claim is rescued from banality by the observation that knowledge itself is multidimensional and that each dimension exerts varying amounts of influence on the numerous aspects of the decision process. Although much work remains before we can achieve a comprehensive mapping of knowledge onto decision making, consumer researchers have warmed to the task. As we have been pursuing the goal, a related area of interest has blossomed—largely over the past decade and largely outside the scope of consumer research—that has been concerned not with the contents of our knowledge but rather with our knowledge of our knowledge. Thus, the dimensions of expertise must be expanded to include the dimension of meta-knowledge.

[*Received July 1999. Revised January 2000. Robert E. Burnkrant served as editor, and Barbara Loken served as associate editor for this article.*]

REFERENCES

Adams, Joe K. and Pauline Austin Adams (1961), "Realism of Confidence in Judgments," *Psychological Review*, 68 (January), 33–45.

Alba, Joseph W. and Amitava Chattopadhyay (1986), "Salience Effects in Brand Recall," *Journal of Marketing Research*, 23 (November), 363–369.

——— and J. Wesley Hutchinson (1987), "Dimensions of Consumer Expertise," *Journal of Consumer Research*, 13 (March), 411–454.

———, J. Wesley Hutchinson, and John G. Lynch, Jr. (1991), "Memory and Decision Making," in *Handbook of Consumer Research*, ed. Thomas S. Robertson and Harold H. Kassarjian, Englewood Cliffs, NJ: Prentice-Hall, 1–49.

Alicke, Mark D. (1985), "Global Self-Evaluation as Determined by the Desirability and Controllability of Trait Adjectives," *Journal of Personality and Social Psychology*, 49 (December), 1621–1630.

Allwood, Carl Martin and Par Anders Granhag (1996), "Realism in Confidence as a Function of Working in Dyads or Alone," *Organizational Behavior and Human Decision Processes*, 66 (June), 277–289.

——— and Henry Montgomery (1987), "Response Selection Strategies and Realism of Confidence Judgments," *Organizational Behavior and Human Decision Processes*, 39 (June), 365–383.

Anderson, Craig A. (1983a), "Abstract and Concrete Data in the Perseverance of Social Theories: When Weak Data Lead to Unshakeable Beliefs," *Journal of Experimental Social Psychology*, 19 (March), 93–108.

——— (1983b), "Imagination and Expectation: The Effect of Imagining Behavioral Scripts on Personal Intentions," *Journal of Personality and Social Psychology*, 45 (August), 293–305.

——— and Sandra S. Godfrey (1987), "Thoughts about Actions: The Effects of Specificity and Availability of Imagined Behavioral Scripts on Expectations about Oneself and Others," *Social Cognition*, 5 (3), 238–258.

———, Mark R. Lepper, and Lee Ross (1980), "Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information," *Journal of Personality and Social Psychology*, 39 (December), 1037–1049.

——— and Elizabeth S. Sechler (1986), "Effects of Explanation and Counterexplanation on the Development and Use of Social Theories," *Journal of Personality and Social Psychology*, 50 (January), 24–34.

Anderson, Urton and William F. Wright (1988), "Expertise and the Explanation Effect," *Organizational Behavior and Human Decision Processes*, 42 (October), 250–269.

Arkes, Hal R., Caryn Christensen, Cheryl Lai, and Catherine Blumer (1987), "Two Methods of Reducing Overconfidence," *Organizational Behavior and Human Decision Processes*, 39 (February), 133–144.

———, Robyn M. Dawes, and Caryn Christensen (1986), "Factors Influencing the Use of a Decision Rule in a Probabilistic Task," *Organizational Behavior and Human Decision Processes*, 37 (February), 93–110.

———, Catherine Hackett, and Larry Boehm (1989), "The Generality of the Relation between Familiarity and Judged Validity," *Journal of Behavioral Decision Making*, 2 (April–June), 81–94.

——— and Allan R. Harkness (1980), "Effect of Making a Decision on Subsequent Recognition of Symptoms," *Journal of Experimental Psychology: Human Learning and Memory*, 6 (September), 568–575.

Armor, David A. and Shelley E. Taylor (1998), "Situated Optimism: Specific Outcome Expectancies and Self-Regulation," in *Advances in Experimental Social Psychology*, Vol. 30, ed. Mark P. Zanna, San Diego, CA: Academic Press, 309–379.

Ayton, Peter and Alastair G. R. McClelland (1997), "How Real

Is Overconfidence?" *Journal of Behavioral Decision Making*, 10 (September), 279–285.

Babad, Elish and Yosi Katz (1991), "Wishful Thinking—Against All Odds," *Journal of Applied Social Psychology*, 21 (December), 1921–1938.

Bacon, Frederick T. (1979), "Credibility of Repeated Statements: Memory for Trivia," *Journal of Experimental Psychology: Human Learning and Memory*, 5 (May), 241–252.

Bahrick, Harry P., Lynda K. Hall, and Stephanie A. Berger (1996), "Accuracy and Distortion in Memory for High School Grades," *Psychological Science*, 7 (September), 265–271.

Barclay, Craig R. and Henry M. Wellman (1986), "Accuracies and Inaccuracies in Autobiographical Memories," *Journal of Memory and Language*, 25 (February), 93–103.

Bargh, John A. and Tanya L. Chartrand (1999), "The Unbearable Automaticity of Being," *American Psychologist*, 54 (July), 462–479.

Begg, Ian Maynard, Ann Anas, and Suzanne Farinacci (1992), "Dissociation of Process in Belief: Source Recollection, Statement Familiarity, and the Illusion of Truth," *Journal of Experimental Psychology: General*, 121 (December), 446–458.

———, Victoria Armour, and Therese Kerr (1985), "On Believing What We Remember," *Canadian Journal of Behavioural Science*, 17 (July), 199–214.

———, Rohan K. Robertson, Vincenza Gruppuso, Ann Anas, and Douglas R. Needham (1996), "The Illusory-Knowledge Effect," *Journal of Memory and Language*, 35 (June), 410–433.

Bell, Brad E. and Elizabeth F. Loftus (1989), "Trivial Persuasion in the Courtroom: The Power of a (Few) Minor Details," *Journal of Personality and Social Psychology*, 56 (May), 669–679.

Benjamin, Aaron S., Robert A. Bjork, and Bennet L. Schwartz (1998), "The Mismeasure of Memory: When Retrieval Fluency Is Misleading," *Journal of Experimental Psychology: General*, 127 (March), 55–68.

Berry, Dianne C. (1996), "How Implicit Is Implicit Learning?" in *Implicit Cognition*, ed. Geoffrey D. M. Underwood, Oxford: Oxford University Press, 203–225.

——— and Zoltan Dienes (1993), *Implicit Learning: Theoretical and Empirical Issues*, Hillsdale, NJ: Erlbaum.

Beyth-Marom, Ruth and Baruch Fischhoff (1983), "Diagnosticity and Pseudodiagnosticity," *Journal of Personality and Social Psychology*, 45 (December), 1185–1195.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1998), "Learning from the Behaviors of Others: Conformity, Fads, and Informational Cascades," *Journal of Economic Perspectives*, 12 (Summer), 151–170.

Bjork, Robert A. (1994), "Memory and Metamemory Considerations in the Training of Human Beings," in *Metacognition: Knowing about Knowing*, ed. Janet Metcalfe and Arthur P. Shimamura, Cambridge, MA: MIT Press, 185–205.

Bornstein, Brian H. and Douglas J. Zickafoose (1999), "'I Know I Know It, I Know I Saw It': The Stability of the Confidence-Accuracy Relationship across Domains," *Journal of Experimental Psychology: Applied*, 5 (March), 76–88.

Bothwell, Robert K., Kenneth A. Deffenbacher, and John C. Brigham (1987), "Correlation of Eyewitness Accuracy and Confidence: Optimality Hypothesis Revisited," *Journal of Applied Psychology*, 72 (November), 691–695.

Bradley, James V. (1981), "Overconfidence in Ignorant Experts," *Bulletin of the Psychonomic Society*, 17 (February), 82–84.

Braun, Phillip A. and Ilan Yaniv (1992), "A Case Study of Expert Judgment: Economists' Probabilities versus Base-Rate Model Forecasts," *Journal of Behavioral Decision Making*, 5 (July–September), 217–231.

Brehmer, Berndt (1980), "In One Word: Not from Experience," *Acta Psychologica*, 45 (August), 223–241.

——— and Jan Kuylenstierna (1978), "Task Information and Performance in Probabilistic Inference Tasks," *Organizational Behavior and Human Performance*, 22 (December), 445–464.

Brenner, Lyle A. (1999), "Should Observed Overconfidence Be Dismissed as a Statistical Artifact? Critique of Erev, Wallsten, & Budescu (1994)," *Psychological Review*, in press.

———, Derek J. Koehler, Varda Liberman, and Amos Tversky (1996), "Overconfidence in Probability and Frequency Judgments: A Critical Examination," *Organizational Behavior and Human Decision Processes*, 65 (March), 212–219.

———, Derek J. Koehler, and Amos Tversky (1996), "On the Evaluation of One-Sided Evidence," *Journal of Behavioral Decision Making*, 9 (March), 59–70.

Broniarczyk, Susan M. and Joseph W. Alba (1994), "The Role of Consumers' Intuitions in Inference Making," *Journal of Consumer Psychology*, 21 (December), 393–407.

Brown, Alan S. and Lori A. Nix (1996), "Turning Lies into Truths: Referential Validation of Falsehoods," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (September), 1088–1100.

Budescu, David V., Ido Erev, and Thomas S. Wallsten (1997), "On the Importance of Random Error in the Study of Probability Judgment. Part I. New Theoretical Developments," *Journal of Behavioral Decision Making*, 10 (September), 157–172.

———, Thomas S. Wallsten, and Wing T. Au (1997), "On the Importance of Random Error in the Study of Probability Judgment. Part II. Applying the Stochastic Judgment Model to Detect Systematic Trends," *Journal of Behavioral Decision Making*, 10 (September), 173–188.

Buehler, Roger, Dale Griffin, and Heather MacDonald (1997), "The Role of Motivated Reasoning in Optimistic Time Predictions," *Personality and Social Psychology Bulletin*, 23 (March), 238–247.

———, Dale Griffin, and Michael Ross (1994), "Exploring the 'Planning Fallacy': Why People Underestimate Their Task Completion Times," *Journal of Personality and Social Psychology*, 67 (September), 366–381.

Byram, Stephanie J. (1997), "Cognitive and Motivational Factors Influencing Time Prediction," *Journal of Experimental Psychology: Applied*, 3 (September), 216–239.

Camerer, Colin and Dan Lovallo (1999), "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 89 (March), 306–318.

Carlson, Bruce W. (1993), "The Accuracy of Future Forecasts and Past Judgments," *Organizational Behavior and Human Decision Processes*, 54 (March), 245–276.

Carroll, John 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 (January), 88–96.

Chandler, C. C. (1994), "Studying Related Pictures Can Reduce Accuracy, but Increase Confidence, in a Modified Recognition Test," *Memory & Cognition*, 22 (May), 273–280.

Cheesman, Jim and Philip M. Merikle (1984), "Priming With and Without Awareness," *Perception and Psychophysics*, 36 (October), 387–395.

——— and Philip M. Merikle (1986), "Distinguishing Conscious from Unconscious Perceptual Processes," *Canadian Journal of Psychology*, 40 (December), 343–367.

Chestnut, Robert W. and Jacob Jacoby (1980), "Product Comprehension: The Case of Permanent vs. Term Life Insurance," in *Advances in Consumer Research*, Vol. 7, ed. Jerry C. Olson, Ann Arbor, MI: Association for Consumer Research, 424–428.

Christenfeld, Nicholas (1995), "Choices from Identical Options," *Psychological Science*, 6 (January), 50–55.

Clark, Steven E. (1997), "A Familiarity-Based Account of Confidence-Accuracy Inversions in Recognition Memory," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23 (January), 232–238.

Conway, Michael and Michael Ross (1984), "Getting What You Want by Revising What You Had," *Journal of Personality and Social Psychology*, 47 (October), 738–748.

Cooper, Arnold, Carolyn Y. Woo, and William C. Dunkelberg (1988), "Entrepreneurs' Perceived Chances for Success," *Journal of Business Venturing*, 3 (Spring), 97–108.

Costermans, Jean, Guy Lories, and Catherine Ansay (1992), "Confidence Level and Feeling of Knowing in Question Answering: The Weight of Inferential Processes," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (January), 142–150.

Dawes, Robyn and Matthew Mulford (1996), "The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in How We Study Judgment?" *Organizational Behavior and Human Decision Processes*, 65 (March), 201–211.

Dellarosa, Denise and Lyle E. Bourne, Jr. (1984), "Decisions and Memory: Differential Retrievability of Consistent and Contradictory Evidence," *Journal of Verbal Learning and Verbal Behavior*, 23 (December), 669–682.

Dhar, Ravi and Itamar Simonson (1992), "The Effect of the Focus of Comparison on Consumer Preferences," *Journal of Marketing Research*, 29 (November), 430–440.

Dienes, Zoltan, Gerry T. M. Altmann, Liam Kwan, and Alastair Goode (1995), "Unconscious Knowledge of Artificial Grammars Is Applied Strategically," *Journal of Experimental Psychology*, 21 (September), 1322–1338.

Ditto, Peter H. and David F. Lopez (1992), "Motivated Skepticism: Use of Differential Decision Criteria for Preferred and Nonpreferred Conclusions," *Journal of Personality and Social Psychology*, 63 (October), 568–584.

———, James A. Scepansky, Geoffrey D. Munro, Anne Marie Apanovitch, and Lisa K. Lockhart (1998), "Motivated Sensitivity to Preference-Inconsistent Information," *Journal of Personality and Social Psychology*, 75 (July), 53–69.

Doherty, Michael E., Randall Chadwick, Hugh Garavan, David Barr, and Clifford R. Mynatt (1996), "On People's Understanding of the Diagnostic Implications of Probabilistic Data," *Memory & Cognition*, 24 (July), 644–654.

———, Clifford R. Mynatt, Ryan D. Tweney, and Michael D. Schiavo (1979), "Pseudodiagnosticity," *Acta Psychologica*, 43 (September), 111–121.

Dougherty, Michael R. P., Charles F. Gettys, and Eve E. Ogden (1999), "MINERVA-DM: A Memory Processes Model for Judgments of Likelihood," *Psychological Review*, 106 (January), 180–209.

Dunlosky, John and Thomas O. Nelson (1997), "Similarity between the Cue for Judgments of Learning (JOL) and the Cue for Test Is Not the Primary Determinant of JOL Accuracy," *Journal of Memory and Language*, 36 (January), 34–49.

Dunning, David, Dale W. Griffin, James D. Milojkovic, and Lee Ross (1990), "The Overconfidence Effect in Social Prediction," *Journal of Personality and Social Psychology*, 58 (April), 568–581.

———, Judith A. Meyerowitz, and Amy D. Holzberg (1989), "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability," *Journal of Personality and Social Psychology*, 57 (December), 1082–1090.

——— and Lisa Beth Stern (1994), "Distinguishing Accurate from Inaccurate Eyewitness Identification via Inquiries about Decision Process," *Journal of Personality and Social Psychology*, 67 (November), 818–835.

Edwards, Kari and Edward E. Smith (1996), "A Disconfirmation Bias in the Evaluation of Arguments," *Journal of Personality and Social Psychology*, 71 (July), 5–24.

Einhorn, Hillel J. and Robin M. Hogarth (1978), "Confidence in Judgment: Persistence in the Illusion of Validity," *Psychological Review*, 85 (September), 395–416.

——— and Robin M. Hogarth (1986), "Judging Probable Cause," *Psychological Bulletin*, 99 (January), 3–19.

Erdfelder, Edgar and Axel Buchner (1998), "Decomposing the Hindsight Bias: A Multinomial Processing Tree Model for Separating Recollection and Reconstruction in Hindsight," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (March), 387–414.

Erev, Ido, Thomas S. Wallsten, and David V. Budescu (1994), "Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes," *Psychological Review*, 101 (July), 519–527.

Feldman, Jack M. and John G. Lynch, Jr. (1988), "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (August), 421–435.

Fischhoff, Baruch (1991), "Value Elicitation: Is There Anything in There?" *American Psychologist*, 46 (August), 835–847.

——— and Don MacGregor (1982), "Subjective Confidence in Forecasts," *Journal of Forecasting*, 1 (April–June), 155–172.

——— and Paul Slovic (1980), "A Little Learning . . . : Confidence in Multicue Judgment Tasks," in *Attention and Performance VIII*, ed. Raymond S. Nickerson, Hillsdale, NJ: Erlbaum, 779–800.

———, Paul Slovic, and Sarah Lichtenstein (1977), "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology: Human Perception and Performance*, 3 (November), 552–564.

———, Paul Slovic, and Sarah Lichtenstein (1978), "Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation," *Journal of Experimental Psychology: Human Perception and Performance*, 4 (May), 330–344.

———, Paul Slovic, and Sarah Lichtenstein (1980), "Knowing What You Want: Measuring Labile Values," in *Cognitive Processes in Choice and Decision Behavior*, ed. Thomas S. Wallsten, Hillsdale, NJ: Erlbaum, 117–141.

Fitzsimons, Gavan J. and Vicki G. Morwitz (1996), "The Effect of Measuring Intent on Brand-Level Purchase Behavior," *Journal of Consumer Research*, 23 (June), 1–11.

Frey, Dieter (1986), "Recent Research on Selective Exposure to Information," in *Advances in Experimental Social Psychology*, Vol. 19, ed. Leonard Berkowitz, Orlando, FL: Academic Press, 41–80.

Ganzach, Yoav and David H. Krantz (1991), "The Psychology of Moderate Prediction. II. Leniency and Uncertainty," *Organizational Behavior and Human Decision Processes*, 48 (April), 169–192.

Gentner, Dedre and Alan Collins (1981), "Studies of Inference from Lack of Knowledge," *Memory & Cognition*, 9 (July), 434–443.

Gettys, Charles F., Thomas Mehle, and Stanley Fisher (1986), "Plausibility Assessments in Hypothesis Generation," *Organizational Behavior and Human Decision Processes*, 37 (February), 14–33.

———, Rebecca M. Pliske, Carol Manning, and Jeff T. Casey (1987), "An Evaluation of Act Generation Performance," *Organizational Behavior and Human Decision Processes*, 39 (February), 23–51.

Gibson, Bryan, David M. Sanbonmatsu, and Steven S. Posavac (1997), "The Effects of Selective Hypothesis Testing on Gambling," *Journal of Experimental Psychology: Applied*, 3 (June), 126–142.

Gigerenzer, Gerd, Ulrich Hoffrage, and Heinz Kleinbolting (1991), "Probabilistic Mental Models: A Brunswikian Theory of Confidence," *Psychological Review*, 98 (October), 506–528.

Gilbert, Daniel T. (1991), "How Mental Systems Believe," *American Psychologist*, 46 (February), 107–119.

——— and Randall E. Osborne (1989), "Thinking Backward: Some Curable and Incurable Consequences of Cognitive Busyness," *Journal of Personality and Social Psychology*, 57 (December), 940–949.

Gilovich, Thomas (1983), "Biased Evaluation and Persistence in Gambling," *Journal of Personality and Social Psychology*, 44 (June), 1110–1126.

——— and Christine Douglas (1986), "Biased Evaluations of Randomly Determined Gambling Outcomes," *Journal of Experimental Social Psychology*, 22 (May), 228–241.

———, Margaret Kerr, and Victoria Husted Medvec (1993), "Effect of Temporal Perspective on Subjective Confidence," *Journal of Personality and Social Psychology*, 64 (April), 552–560.

Glazer, Rashi, Joel H. Steckel, and Russell S. Winer (1992), "Locally Rational Decision Making: The Distracting Effect of Information on Managerial Performance," *Management Science*, 38 (February), 212–226.

Glenberg, Arthur M., Alex Cherry Wilkinson, and William Epstein (1982), "The Illusion of Knowing: Failure in the Self-Assessment of Comprehension," *Memory & Cognition*, 10 (November), 597–602.

——— and William Epstein (1985), "Calibration of Comprehension," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11 (October), 702–718.

——— and William Epstein (1987), "Inexpert Calibration of Comprehension," *Memory & Cognition*, 15 (January), 84–93.

———, Thomas Sanocki, William Epstein, and Craig Morris (1987), "Enhancing Calibration of Comprehension," *Journal of Experimental Psychology: General*, 116 (June), 119–136.

Gollwitzer, Peter M. and Ronald F. Kinney (1989), "Effects of Deliberative Mind-Sets on Illusion of Control," *Journal of Personality and Social Psychology*, 56 (April), 531–542.

Gregory, W. Larry, Robert B. Cialdini, and Kathleen M. Carpenter (1982), "Self-Relevant Scenarios as Mediators of Likelihood Estimates and Compliance: Does Imagining Make It So?" *Journal of Personality and Social Psychology*, 43 (July), 89–99.

Griffin, Dale and Amos Tversky (1992), "The Weighting of Evidence and the Determinants of Confidence," *Cognitive Psychology*, 24 (July), 411–435.

——— and Carol A. Varey (1996), "Commentary: Toward a Consensus on Overconfidence," *Organizational Behavior and Human Decision Processes*, 65 (March), 227–231.

Harris, Peter (1996), "Sufficient Grounds for Optimism? The Relationship between Perceived Controllability and Optimistic Bias," *Journal of Social and Clinical Psychology*, 15 (Spring), 9–52.

Harvey, Nigel (1994), "Relations between Confidence and Skilled Performance," in *Subjective Probability*, ed. George Wright and Peter Ayton, Chichester: Wiley, 321–352.

———, Derek J. Koehler, and Peter Ayton (1997), "Judgments of Decision Effectiveness: Actor-Observer Differences in Overconfidence," *Organizational Behavior and Human Decision Processes*, 70 (June), 267–282.

Hasher, Lynn, David Goldstein, and Thomas Toppino (1977), "Frequency and the Conference of Referential Validity," *Journal of Verbal Learning and Verbal Behavior*, 16 (February), 107–112.

Hawkins, Scott A. and Reid Hastie (1990), "Hindsight: Biased Judgments of Past Events after the Outcomes Are Known," *Psychological Bulletin*, 107 (May), 311–327.

——— and Stephen J. Hoch (1992), "Low-Involvement Learning: Memory without Evaluation," *Journal of Consumer Research*, 19 (September), 212–225.

Heath, Chip and Rich Gonzalez (1995), "Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence against Information Collection Views of Interactive Decision Making," *Organizational Behavior and Human Decision Processes*, 61 (March), 305–326.

——— and Amos Tversky (1991), "Preference and Belief: Ambiguity and Competence in Choice under Uncertainty," *Journal of Risk and Uncertainty*, 4 (1), 5–28.

Heine, Steven J. and Darrin R. Lehman (1995), "Cultural Variation in Unrealistic Optimism: Does the West Feel More Invulnerable than the East?" *Journal of Personality and Social Psychology*, 68 (April), 595–607.

Henry, Rebecca A. (1994), "The Effects of Choice and Incentives on the Overestimation of Future Performance," *Organizational Behavior and Human Decision Processes*, 57 (February), 210–225.

Hintzman, Douglas L. (1988), "Judgments of Frequency and Recognition Memory in a Multiple-Trace Memory Model," *Psychological Review*, 95 (October), 528–551.

Hirt, Edward R. and Keith D. Markman (1995), "Multiple Explanations: A Consider-an-Alternative Strategy for Debiasing Judgments," *Journal of Personality and Social Psychology*, 69 (December), 1069–1086.

——— and Steven J. Sherman (1985), "The Role of Prior Knowledge in Explaining Hypothetical Events," *Journal of Experimental Social Psychology*, 21 (November), 519–543.

Hoch, Stephen J. (1984), "Availability and Interference in Predictive Judgment," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10 (October), 649–662.

——— (1985), "Counterfactual Reasoning and Accuracy in Predicting Personal Events,"*Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11 (October), 719–731.

——— (1987), "Perceived Consensus and Predictive Accuracy: The Pros and Cons of Projection," *Journal of Personality and Social Psychology*, 53 (August), 221–234.

——— (1988), "Who Do We Know: Predicting the Interests and Opinions of the American Consumer," *Journal of Consumer Research*, 15 (December), 315–324.

——— and John Deighton (1989), "Managing What Consumers

Learn from Experience," *Journal of Marketing*, 53 (April), 1–20.

Holmes, Jonathan B., Harriet Salatas Waters, and Suparna Rajaram (1998), "The Phenomenology of False Memories: Episodic Content and Confidence," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (July), 1026–1040.

Holyoak, Keith J. and Dan Simon (1999), "Bidirectional Reasoning in Decision Making by Constraint Satisfaction," *Journal of Experimental Psychology: General*, 128 (March), 3–31.

Hutchinson, J. Wesley and Joseph W. Alba (1991), "Ignoring Irrelevant Information: Situational Determinants of Consumer Learning," *Journal of Consumer Behavior*, 18 (December), 325–345.

——— and Robert J. Meyer (1994), "Dynamic Decision Making: Optimal Policies and Actual Behavior in Sequential Choice Problems," *Marketing Letters*, 5 (4), 369–382.

Hyman, Ray (1989), "The Psychology of Deception," *Annual Review of Psychology*, 40, 133–154.

Jacoby, Larry L., Colleen Kelley, Judith Brown, and Jennifer Jasechko (1989), "Becoming Famous Overnight: Limits on the Ability to Avoid Unconscious Influences of the Past," *Journal of Personality and Social Psychology*, 56 (March), 326–338.

———, Colleen Kelley, and Jane Dywan (1989), "Memory Attributions," in *Varieties of Memory and Consciousness*, ed. Henry L. Roediger III and Fergus I. M. Craik, Hillsdale, NJ: Erlbaum, 391–422.

Jameson, Anthony, Thomas O. Nelson, R. Jacob Leonesio, and Louis Narens (1993), "The Feeling of Another Person's Knowing," *Journal of Memory and Language*, 32 (June), 320–335.

John, Deborah Roedder, Carol A. Scott, and James R. Bettman (1986), "Sampling Data for Covariation Assessment: The Effects of Prior Beliefs on Search Patterns," *Journal of Consumer Research*, 13 (June), 38–47.

Johnson, Hollyn M. and Colleen M. Seifert (1994), "Sources of the Continued Influence Effect: When Misinformation in Memory Affects Later Inferences," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20 (November), 1420–1436.

Johnson, Marcia, Shahin Hashtroudi, and D. Stephen Lindsay (1993), "Source Monitoring," *Psychological Bulletin*, 114 (July), 3–28.

Jones, Steven K., Kristine Taylor Jones, and Deborah Frisch (1995), "Biases of Probability Assessment: A Comparison of Frequency and Single-Case Judgments," *Organizational Behavior and Human Decision Processes*, 61 (February), 109–122.

———, Tricia J. Yurak, and Deborah Frisch (1997), "The Effect of Outcome Information on the Evaluation and Recall of Individuals' Own Decisions," *Organizational Behavior and Human Decision Processes*, 71 (July), 95–120.

Jungermann, Helmut and Manfre Thuring (1987), "The Use of Mental Models for Generating Scenarios," in *Judgmental Forecasting*, ed. George Wright and Peter Ayton, Chichester: Wiley, 245–266.

Juslin, Peter and Henrik Olsson (1997), "Thurstonian and Brunswikian Origins of Uncertainty in Judgment: A Sampling Model of Confidence in Sensory Discrimination," *Psychological Review*, 104 (April), 344–366.

———, Henrik Olsson, and Mats Bjorkman (1997), "Brunswikian and Thurstonian Origins of Bias in Probability Assessment: On the Interpretation of Stochastic Components of Judgment," *Journal of Behavioral Decision Making*, 10 (September), 189–210.

———, Henrik Olsson, and Anders Winman (1996), "Calibration and Diagnosticity of Confidence in Eyewitness Identification: Comments on What Can Be Inferred from the Low Confidence-Accuracy Correlation," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (September), 1304–1316.

Kahneman, Daniel and Dan Lovallo (1993), "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking," *Management Science*, 39 (January), 17–31.

——— and Jackie Snell (1992), "Predicting a Changing Taste: Do People Know What They Will Like?" *Journal of Behavioral Decision Making*, 5 (July–September), 187–200.

——— and Amos Tversky (1982), "The Simulation Heuristic," in *Judgment under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman, Paul Slovic, and Amos Tversky, Cambridge: Cambridge University Press, 201–208.

Kelemen, William L. and Charles A. Weaver III (1997), "Enhanced Metamemory at Delays: Why Do Judgments of Learning Improve over Time?" *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23 (November), 1394–1409.

Kelley, Colleen M. and D. Stephen Lindsay (1993), "Remembering Mistaken for Knowing: Ease of Retrieval as a Basis for Confidence in Answers to General Knowledge Questions," *Journal of Memory and Language*, 32 (February), 1–24.

Keren, Gideon (1991), "Calibration and Probability Judgments: Conceptual and Methodological Issues," *Acta Psychologica*, 77 (September), 217–273.

——— (1997), "On the Calibration of Probability Judgments: Some Critical Comments and Alternative Perspectives," *Journal of Behavioral Decision Making*, 10 (September), 269–278.

Klayman, Joshua (1995), "Varieties of Confirmation Bias," *The Psychology of Learning and Motivation*, 32, 385–417.

———, Jack B. Soll, Claudia Gonzalez-Vallejo, and Sema Barlas (1999), "Overconfidence: It Depends on How, What, and Who You Ask," *Organizational Behavior and Human Decision Processes*, 79 (September), 216–247.

Kleinmuntz, Benjamin (1990), "Why We Use Our Heads Instead of Formulas: Toward an Integrative Approach," *Psychological Bulletin*, 107 (May), 296–310.

Koehler, Derek J. (1991), "Explanation, Imagination, and Confidence in Judgment," *Psychological Bulletin*, 110 (November), 499–519.

——— (1994), "Hypothesis Generation and Confidence in Judgment," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20 (March), 461–469.

———, Lyle A. Brenner, Varda Liberman, and Amos Tversky (1996), "Confidence and Accuracy in Trait Inference: Judgment by Similarity," *Acta Psychologica*, 92 (June), 33–57.

——— and Nigel Harvey (1997), "Confidence Judgments by Actors and Observers," *Journal of Behavioral Decision Making*, 10 (September), 221–242.

Koehler, Jonathan J. (1993), "The Influence of Prior Beliefs on Scientific Judgments of Evidence Quality," *Organizational Behavior and Human Decision Processes*, 56 (October), 28–55.

Koriat, Asher (1993), "How Do We Know That We Know? The Accessibility Model of the Feeling of Knowing," *Psychological Review*, 100 (October), 609–639.

———, Sarah Lichtenstein, and Baruch Fischhoff (1980), "Rea-

sons for Confidence," *Journal of Experimental Psychology: Human Learning and Memory*, 6 (March), 107–118.

Kottemann, Jeffrey E., Fred D. Davis, and William E. Remus (1994), "Computer-Assisted Decision Making: Performance, Beliefs, and the Illusion of Control," *Organizational Behavior and Human Decision Processes*, 57 (January), 26–37.

Krueger, Joachim (1998), "On the Perception of Social Consensus," in *Advances in Experimental Social Psychology*, Vol. 30, ed. Mark P. Zanna, San Diego, CA: Academic Press, 163–240.

Kruger, Justin and David Dunning (1999), "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments," *Journal of Personality and Social Psychology*, 77 (December), 1121–1134.

Kruglanski, Arie W. and Donna M. Webster (1996), "Motivated Closing of the Mind: 'Seizing' and 'Freezing,'" *Psychological Review*, 103 (April), 263–283.

Kuhn, Deanna, Michael Weinstock, and Robin Flaton (1994), "How Well Do Jurors Reason? Competence Dimensions of Individual Variation in a Juror Reasoning Task," *Psychological Science*, 5 (September), 289–296.

Kuhn, Kristine M. and Janet A. Sniezek (1996), "Confidence and Uncertainty in Judgmental Forecasting: Differential Effects of Scenario Presentation," *Journal of Behavioral Decision Making*, 9 (December), 231–247.

Kunda, Ziva (1987), "Motivated Inference: Self-Serving Generation and Evaluation of Causal Theories," *Journal of Personality and Social Psychology*, 53 (October), 636–647.

——— (1990), "The Case for Motivated Reasoning," *Psychological Bulletin*, 108 (November), 480–498.

——— and Kathryn C. Oleson (1997), "When Exceptions Prove the Rule: How Extremity of Deviance Determines the Impact of Deviant Examples on Stereotypes," *Journal of Personality and Social Psychology*, 72 (May), 965–979.

Langer, Ellen J. (1975), "The Illusion of Control," *Journal of Personality and Social Psychology*, 32 (August), 311–328.

Larwood, Laurie and William Whittaker (1977), "Managerial Myopia: Self-Serving Biases in Organizational Planning," *Journal of Applied Psychology*, 62 (April), 194–198.

Law, Sharmistha and Scott A. Hawkins (1997), "Advertising Repetition and Consumer Beliefs: The Role of Source Memory," in *Measuring Advertising Effectiveness*, ed. William D. Wells, Mahwah, NJ: Erlbaum, 67–75.

Lerner, Jennifer S. and Philip E. Tetlock (1999), "Accounting for the Effects of Accountability," *Psychological Bulletin*, 125 (March), 255–275.

Levi, Ariel S. and John B. Pryor (1987), "Use of the Availability Heuristic in Probability Estimates of Future Events: The Effects of Imagining Outcomes versus Imagining Reasons," *Organizational Behavior and Human Decision Processes*, 40 (October), 219–234.

Levin, Irwin P., Daniel P. Chapman, and Richard D. Johnson (1988), "Confidence in Judgments Based on Incomplete Information: An Investigation Using Both Hypothetical and Real Gambles," *Journal of Behavioral Decision Making*, 1 (January–March), 29–41.

Levine, Gary M., Jamin B. Halberstadt, and Robert L. Goldstone (1996), "Reasoning and the Weighting of Attributes in Attitude Judgments," *Journal of Personality and Social Psychology*, 70 (February), 230–240.

Levine, Linda J. (1997), "Reconstructing Memory for Emotions," *Journal of Experimental Psychology: General*, 126 (June), 165–177.

Liberman, Varda and Amos Tversky (1993), "On the Evaluation of Probability Judgments: Calibration, Resolution, and Monotonicity," *Psychological Bulletin*, 114 (July), 162–173.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D. Phillips (1982), "Calibration of Probabilities: The State of the Art to 1980," in *Judgment under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman, Paul Slovic, and Amos Tversky, Cambridge: Cambridge University Press, 306–334.

Linville, Patricia W. (1982), "The Complexity-Extremity Effect and Age-Based Stereotyping," *Journal of Personality and Social Psychology*, 42 (February), 193–211.

Loftus, Elizabeth F., Karen Donders, Hunter G. Hoffman, and Jonathan W. Schooler (1989), "Creating New Memories That Are Quickly Accessed and Confidently Held," *Memory & Cognition*, 17 (September), 607–616.

Luce, Mary Frances (1998), "Choosing to Avoid: Coping with Negatively Emotion-Laden Consumer Decisions," *Journal of Consumer Research*, 24 (March), 409–433.

Lusk, Cynthia M. and Charles M. Judd (1988), "Political Expertise and the Structural Mediators of Candidate Evaluations," *Journal of Experimental Social Psychology*, 24 (March), 105–126.

Luus, C. A. Elizabeth and Gary L. Wells (1994), "The Malleability of Eyewitness Confidence: Co-Witness and Perseverance Effects," *Journal of Applied Psychology*, 79 (October), 714–723.

Mabe, Paul A. III and Stephen G. West (1982), "Validity of Self-Evaluation of Ability: A Review and Meta-Analysis," *Journal of Applied Psychology*, 67 (June), 280–296.

Mahajan, Jayashree (1992), "The Overconfidence Effect in Marketing Management Predictions," *Journal of Marketing Research*, 29 (August), 329–342.

Maki, Ruth H. (1995), "Accuracy of Metacomprehension Judgments for Questions of Varying Importance Levels," *American Journal of Psychology*, 108 (Fall), 327–344.

——— (1998a), "Predicting Performance on Text: Delayed versus Immediate Predictions and Tests," *Memory & Cognition*, 26 (September), 959–964.

——— (1998b), "Metacomprehension of Text: Influence of Absolute Confidence Level on Bias and Accuracy," in *The Psychology of Learning and Memory*, Vol. 38, ed. Douglas L. Medin, San Diego, CA: Academic Press, 223–248.

———, James M. Foley, Wendy K. Kajer, Richard C. Thompson, and Meryl G. Willert (1990), "Increased Processing Enhances Calibration of Comprehension," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16 (July), 609–616.

——— and Matt Serra (1992), "The Basis of Test Predictions for Text Material," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (January), 116–126.

Marks, Gary and Norman Miller (1987), "Ten Years of Research on the False-Consensus Effect: An Empirical and Theoretical Review," *Psychological Bulletin*, 102 (July), 72–90.

Mayesless, Ofra and Arie W. Kruglanski (1987), "What Makes You So Sure? Effects of Epistemic Motivations on Judgmental Confidence," *Organizational Behavior and Human Decision Processes*, 39 (April), 162–183.

Mazumdar, Tridib and Kent B. Monroe (1992), "Effects of Inter-Store and In-Store Price Comparisons on Price Recall Accuracy and Confidence," *Journal of Retailing*, 68 (Spring), 66–89.

Mazzoni, Giuliana and Thomas O. Nelson (1995), "Judgments of Learning Are Affected by the Kind of Encoding in Ways That Cannot Be Attributed to the Level of Recall," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21 (September), 1263–1274.

McClelland, Alastair G. R. and Fergus Bolger (1994), "The Calibration of Subjective Probabilities: Theories and Models 1980–94," in *Subjective Probability*, ed. George Wright and Peter Ayton, Chichester: Wiley, 453–482.

Mehle, Thomas (1982), "Hypothesis Generation in an Automobile Malfunction Inference Task," *Acta Psychologica*, 52 (November), 87–106.

———, Charles F. Gettys, Carol Manning, Suzanne Baca, and Stanley Fisher (1981), "The Availability Explanation of Excessive Plausibility Assessments," *Acta Psychologica*, 49 (November), 127–140.

Metcalfe, Janet (1998), "Cognitive Optimism: Self-Deception or Memory-Based Processing Heuristics," *Personality and Social Psychology Review*, 2 (2), 100–110.

Milburn, Michael (1978), "Sources of Bias in the Prediction of Future Events," *Organizational Behavior and Human Performance*, 21 (February), 17–26.

Millar, Murray G. and Abraham Tesser (1986), "Thought-Induced Attitude Change: The Effects of Schema Structure and Commitment," *Journal of Personality and Social Psychology*, 51 (August), 259–269.

Moeckel, Cindy L. and R. David Plumlee (1989), "Auditors' Confidence in Recognition of Audit Evidence," *Accounting Review*, 64 (October), 653–666.

Moorman, Christine (1999), "Knowledge Illusions and the Functionality of Overconfidence," paper presented at the annual conference of the Association for Consumer Research, Columbus, OH.

Morris, Carl Craig (1990), "Retrieval Processes Underlying Confidence in Comprehension Judgments," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16 (March), 223–232.

Muthukrishnan, A. V. (1995), "Decision Ambiguity and Incumbent Brand Advantage," *Journal of Consumer Research*, 22 (June), 98–109.

Mynatt, Clifford R., Michael E. Doherty, and William Dragan (1993), "Information Relevance, Working Memory, and the Consideration of Alternatives," *Quarterly Journal of Experimental Psychology*, 46A (November), 759–778.

Nedungadi, Prakash (1990), "Recall and Consumer Consideration Sets: Influencing Choice without Altering Brand Evaluations," *Journal of Consumer Research*, 17 (December), 263–276.

Neisser, Ulric (1991), "A Case of Misplaced Nostalgia," *American Psychologist*, 46 (January), 34–36.

Nelson, Thomas O. and John Dunlosky (1991), "When People's Judgments of Learning (JOLs) Are Extremely Accurate at Predicting Subsequent Recall: The 'Delayed-JOL Effect,'" *Psychological Science*, 2 (July), 267–270

Nisbett, Richard E. and Timothy DeCamp Wilson (1977a), "The Halo Effect: Evidence for Unconscious Alteration of Judgments," *Journal of Personality and Social Psychology*, 35 (April), 250–256.

——— and Timothy DeCamp Wilson (1977b), "Telling More than We Can Know: Verbal Reports on Mental Processes," *Psychological Review*, 84 (May), 231–259.

O'Connor, Marcus (1989), "Models of Human Behavior and Confidence in Judgment: A Review," *International Journal of Forecasting*, 5 (2), 159–169.

Olsen, Robert A. (1997), "Desirability Bias among Professional Investment Managers: Some Evidence from Experts," *Journal of Behavioral Decision Making*, 10 (March), 65–72.

Olsson, Nils, Peter Juslin, and Anders Winman (1998), "Realism of Confidence in Earwitness versus Eyewitness Identification," *Journal of Experimental Psychology: Applied*, 4 (June), 101–118.

Oskamp, Stuart (1965), "Overconfidence in Case-Study Judgments," *Journal of Consulting Psychology*, 29 (June), 261–265.

Paese, Paul W. and Michael A. Feuer (1991), "Decisions, Actions, and the Appropriateness of Confidence in Knowledge," *Journal of Behavioral Decision Making*," 4 (January–March), 1–16.

——— and Janet A. Sniezek (1991), "Influences on the Appropriateness of Confidence in Judgment: Practice, Effort, Information, and Decision-Making," *Organizational Behavior and Human Decision Processes*, 48 (February), 100–130.

Park, C. Whan, David L. Mothersbaugh, and Lawrence Feick (1994), "Consumer Knowledge Assessment," *Journal of Consumer Research*, 21 (June), 71–82.

Payne, John W., James R. Bettman, and Eric J. Johnson (1988), "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14 (July), 534–552.

Pennington, Nancy and Reid Hastie (1988), "Explanation-Based Decision Making: Effects of Memory Structure on Judgment," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14 (July), 521–533.

Perfect, Timothy J. and Tara S. Hollins (1999), "Feeling-of-Knowing Judgments Do Not Predict Subsequent Recognition Performance for Eyewitness Memory," *Journal of Experimental Psychology: Applied*, 5 (September), 250–264.

Perloff, Linda S. and Barbara K. Fetzer (1986), "Self-Other Judgments and Perceived Vulnerability to Victimization," *Journal of Personality and Social Psychology*, 50 (March), 502–510.

Peterson, Dane K. and Gordon F. Pitz (1986), "Effects of Amount of Information on Predictions of Uncertain Quantities," *Acta Psychologica*, 61 (April), 229–241.

——— and Gordon F. Pitz (1988), "Confidence, Uncertainty, and the Use of Information," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14 (January), 85–92.

Posavac, Steven S., David M. Sonbonmatsu, and Russell H. Fazio (1997), "Considering the Best Choice: Effects of the Salience and Accessibility of Alternatives on Attitude-Decision Consistency," *Journal of Personality and Social Psychology*, 72 (February), 253–261.

Powell, Jack L. (1991), "An Attempt at Increasing Decision Rule Usage in a Judgment Task," *Organizational Behavior and Human Decision Processes*, 48 (February), 89–99.

Pratkanis, Anthony R., Anthony G. Greenwald, Michael R. Leippe, and Michael H. Baumgardner (1988), "In Search of Reliable Persuasion Effects. III. The Sleeper Effect Is Dead. Long Live the Sleeper Effect," *Journal of Personality and Social Psychology*, 54 (February), 203–218.

Pritchard, Mary E. and Janice M. Keenan (1999), "Memory Monitoring in Mock Jurors," *Journal of Experimental Psychology: Applied*, 5 (June), 152–168.

Pulford, Briony D. and Andrew M. Colman (1996), "Overconfidence, Base Rates, and Outcome Positivity/Negativity of Predicted Outcomes," *British Journal of Psychology*, 87 (August), 431–445.

Quadrel, Marilyn Jacobs, Baruch Fischhoff, and Wendy Davis (1993), "Adolescent (In)vulnerability," *American Psychologist*, 48 (February), 102–116.

Radecki, Carmen M. and James Jaccard (1995), "Perceptions of Knowledge, Actual Knowledge, and Information Search Be-

havior," *Journal of Experimental Social Psychology*, 31 (March), 107–138.

Reder, Lynne M. and Frank E. Ritter (1992), "What Determines Initial Feelings of Knowing? Familiarity with Question Terms, Not with the Answer," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (May), 435–451.

Ricchiute, David N. (1998), "Evidence, Memory, and Causal Order in a Complex Audit Decision Task," *Journal of Experimental Psychology: Applied*, 4 (March), 3–15.

Roediger, Henry L. III and Kathleen B. McDermott (1995), "Creating False Memories: Remembering Words Not Presented in Lists," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21 (July), 803–814.

Ronis, David L. and J. Frank Yates (1987), "Components of Probability Judgment Accuracy: Individual Consistency and Effects of Subject Matter and Assessment Method," *Organizational Behavior and Human Decision Processes*, 40 (October), 193–218.

Ross, Michael (1989), "Relation of Implicit Theories to the Construction of Personal Histories," *Psychological Review*, 96 (April), 341–357.

Russo, J. Edward and Karen J. Kolzow (1994), "Where Is the Fault in Fault Trees?" *Journal of Experimental Psychology: Human Perception and Performance*, 20 (February), 17–32.

———, Victoria Husted Medvec, and Margaret G. Meloy (1996), "The Distortion of Information during Decisions," *Organizational Behavior and Human Decision Processes*, 66 (April), 102–110.

———, Margaret G. Meloy, and Victoria Husted Medvec (1998), "Predecisional Distortion of Product Information," *Journal of Marketing Research*, 35 (November), 438–452.

——— and Paul J. H. Schoemaker (1989), *Decision Traps: The Ten Barriers to Brilliant Decision-Making and How to Overcome Them*, New York: Simon & Schuster.

——— and Paul J. H. Schoemaker (1992), "Managing Overconfidence," *Sloan Management Review*, 33 (Winter), 7–17.

Sanbonmatsu, David M., Sharon A. Akimoto, and Earlene Biggs (1993), "Overestimating Causality: Attributional Effects of Confirmatory Processing," *Journal of Personality and Social Psychology*, 65 (November), 892–903.

———, Frank R. Kardes, and Paul M. Herr (1992), "The Role of Prior Knowledge and Missing Information in Multiattribute Evaluation," *Organizational Behavior and Human Decision Processes*, 51 (February), 76–91.

———, Frank R. Kardes, and Carol Sansone (1991), "Remembering Less and Inferring More: Effects of Time of Judgment on Inference about Unknown Attributes," *Journal of Personality and Social Psychology*, 61 (October), 546–554.

———, Steven S. Posavac, Frank R. Kardes, and Susan P. Mantel (1998), "Selective Hypothesis Testing," *Psychonomic Bulletin and Review*, 5 (June), 197–220.

Sanna, Lawrence J. (1999), "Mental Simulations, Affect, and Subjective Confidence: Timing Is Everything," *Psychological Science*, 10 (July), 339–345.

Santioso, Rasyid and Ziva Kunda (1991), "Ducking the Collection of Costly Evidence: Motivated Use of Statistical Heuristics," *Journal of Behavioral Decision Making*, 4 (July–September), 161–178.

Schacter, Daniel L. (1987), "Implicit Memory: History and Current Status," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13 (July), 501–518

Schneider, Sandra L. (1995), "Item Difficulty, Discrimination, and the Confidence-Frequency Effect in a Categorical Judgment Task," *Organizational Behavior and Human Decision Processes*, 61 (February), 148–167.

——— and Suzanne K. Laurion (1993), "Do We Know What We've Learned from Listening to the News?" *Memory & Cognition*, 21 (March), 198–209.

Schwartz, Bennet L. and Janet Metcalfe (1992), "Cue Familiarity but Not Target Retrievability Enhances Feeling-of-Knowing Judgments," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (September), 1074–1083.

Sedikides, Constantine and Michael J. Strube (1997), "Self-Evaluation: To Thine Own Self Be Good, To Thine Own Self Be Sure, To Thine Own Self Be True, and to Thine Own Self Be Better," in *Advances in Experimental Social Psychology*, Vol. 29, ed. Mark P. Zanna, San Diego, CA: Academic Press, 209–269.

Seger, Carol Augart (1994), "Implicit Learning," *Psychological Bulletin*, 115 (March), 163–196.

Shafir, Eldar and Amos Tversky (1992), "Thinking through Uncertainty: Nonconsequential Reasoning and Choice," *Cognitive Psychology*, 24 (October), 449–474.

Shaklee, Harriet and Baruch Fischhoff (1982), "Strategies of Information Search in Causal Analysis," *Memory & Cognition*, 10 (November), 520–530.

Sharp, Glen L., Brian L. Cutler, and Steven D. Penrod (1988), "Performance Feedback Improves the Resolution of Confidence Judgments," *Organizational Behavior and Human Decision Processes*, 42 (December), 271–283.

Shaw, John S. III (1996), "Increases in Eyewitness Confidence Resulting from Postevent Questioning," *Journal of Experimental Psychology: Applied*, 2 (June), 126–146.

Shepperd, James A., Judith A. Ouellette, and Julie K. Fernandez (1996), "Abandoning Unrealistic Optimism: Performance Estimates and the Temporal Proximity of Self-Relevant Feedback," *Journal of Personality and Social Psychology*, 70 (April), 844–855.

Sherman, Steven J., Robert B. Cialdini, Donna F. Schwartzman, and Kim D. Reynolds (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 (March), 118–127.

———, Richard B. Skov, Esther F. Hervitz, and Caryl B. Stock (1981), "The Effects of Explaining Hypothetical Future Events: From Possibility to Probability to Actuality and Beyond," *Journal of Experimental Social Psychology*, 17 (March), 142–158.

Sieck, Winston and J. Frank Yates (1997), "Exposition Effects on Decision Making: Choice and Confidence in Choice," *Organizational Behavior and Human Decision Processes*, 70 (June), 207–219.

Simmons, Carolyn J., Barbara A. Bickart, and John G. Lynch, Jr. (1993), "Capturing and Creating Public Opinion in Survey Research," *Journal of Consumer Research*, 20 (September), 316–329.

Simonson, Itamar (1990), "The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior," *Journal of Marketing Research*, 27 (May), 150–162.

Skov, Richard B. and Steven J. Sherman (1986), "Information-Gathering Processes: Diagnosticity, Hypothesis-Confirming Strategies, and Perceived Hypothesis Confirmation," *Journal of Experimental Social Psychology*, 22 (March), 93–121.

Slusher, Morgan P. and Craig A. Anderson (1987), "When Reality Monitoring Fails: The Role of Imagination in Stereotype

Maintenance," *Journal of Personality and Social Psychology*, 52 (April), 653–662.

Sniezek, Janet A. (1992), "Groups under Uncertainty: An Examination of Confidence in Group Decision Making," *Organizational Behavior and Human Decision Processes*, 52 (June), 124–155.

——— and Timothy Buckley (1991), "Confidence Depends on Level of Aggregation," *Journal of Behavioral Decision Making*, 4 (October–December), 263–272.

———, Paul W. Paese, and Fred S. Switzer III (1990), "The Effect of Choosing on Confidence in Choice," *Organizational Behavior and Human Decision Processes*, 46 (August), 264–282.

Soll, Jack B. (1996), "Determinants of Overconfidence and Miscalibration: The Roles of Random Error and Ecological Structure," *Organizational Behavior and Human Decision Processes*, 65 (February), 117–137.

Spence, Mark T. (1996), "Problem–Problem Solver Characteristics Affecting the Calibration of Judgments," *Organizational Behavior and Human Decision Processes*, 67 (September), 271–279.

Sporer, Siegfried Ludwig, Steven Penrod, Don Read, and Brian Cutler (1995), "Choosing, Confidence, and Accuracy: A Meta-Analysis of the Confidence-Accuracy Relation in Eyewitness Identification Studies," *Psychological Bulletin*, 118 (November), 315–327.

Stone, Dan N. (1994), "Overconfidence in Initial Self-Efficacy Judgments: Effects on Decision Processes and Performance," *Organizational Behavior and Human Decision Processes*, 59 (September), 452–474.

Suantak, Liana, Fergus Bolger, and William Ferrell (1996), "The Hard-Easy Effect in Subjective Probability Calibration," *Organizational Behavior and Human Decision Processes*, 67 (August), 201–221.

Svenson, Ola (1981), "Are We All Less Risky and More Skillful than Our Fellow Drivers?" *Acta Psychologica*, 47 (February), 143–148.

——— (1992), "Differentiation and Consolidation Theory of Human Decision Making: A Frame of Reference for the Study of Pre- and Post-Decision Processes," *Acta Psychologica*, 80 (August), 143–168.

——— and Lars J. Benthorn (1992), "Consolidation Processes in Decision Making: Post-Decision Changes in Attractiveness of Alternatives," *Journal of Economic Psychology*, 13 (June), 315–327.

Swann, William B. Jr. and Michael J. Gill (1997), "Confidence and Accuracy in Person Perception: Do We Know What We Think We Know about Our Relationship Partners?" *Journal of Personality and Social Psychology*, 73 (October), 747–757.

Taylor, Shelley E. and Jonathon D. Brown (1988), "Illusion and Well-Being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin*, 103 (March), 193–210.

———, Lien B. Pham, Inna D. Rivkin, and David A. Armor (1998), "Harnessing the Imagination: Mental Simulation, Self-Regulation, and Coping," *American Psychologist*, 53 (April), 429–439.

Tesser, Abraham (1978), "Self-Generated Attitude Change," in *Advances in Experimental Social Psychology*, Vol. 11, ed. Leonard Berkowitz, New York: Academic Press, 289–338.

Tetlock, Philip E. and Jae Il Kim (1987), "Accountability and Judgment in a Personality Prediction Task," *Journal of Personality and Social Psychology*, 52 (April), 700–709.

Trafimow, David and Janet A. Sniezek (1994), "Perceived Expertise and Its Effect on Confidence," *Organizational Behavior and Human Decision Processes*, 57 (February), 290–302.

Treadwell, Jonathan R. and Thomas O. Nelson (1996), "Availability of Information and the Aggregation of Confidence in Prior Decisions," *Organizational Behavior and Human Decision Processes*, 68 (October), 13–27.

Tschirgi, Judith E. (1980), "Sensible Reasoning: A Hypothesis about Hypotheses," *Child Development*, 51 (March), 1–10.

Tulving, Endel (1983), *Elements of Episodic Memory*, New York: Oxford University Press.

Tversky, Amos and Daniel Kahneman (1973), "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 5 (March), 207–232.

——— and Daniel Kahneman (1983), "Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *Psychological Review*, 90 (October), 293–315.

——— and Derek J. Koehler (1994), "Support Theory: A Nonextensional Representation of Subjective Probability," *Psychological Review*, 101 (October), 547–567.

Underwood, Geoffrey and James E. H. Bright, (1996), "Cognition With and Without Awareness," in *Implicit Cognition*, ed. Geoffrey D. M. Underwood, Oxford: Oxford University Press, 1–40.

Vallone, Robert P., Dale W. Griffin, Sabrina Lin, and Lee Ross (1990), "Overconfident Prediction of Future Actions and Outcomes by Self and Others," *Journal of Personality and Social Psychology*, 58 (April), 582–592.

Van Wallendael, Lori Robinson, and Reid Hastie (1990), "Tracing the Footsteps of Sherlock Holmes: Cognitive Representations of Hypothesis Testing," *Memory & Cognition*, 18 (May), 240–250.

Wagenaar, Willem A. (1988), "Calibration and the Effects of Knowledge and Reconstruction in Retrieval from Memory," *Cognition*, 28 (April), 277–296.

Wallsten, Thomas S. (1996), "Commentary: An Analysis of Judgment Research Analyses," *Organizational Behavior and Human Decision Processes*, 65 (March), 220–226.

———, Randall H. Bender, and Yuenlin Li (1999), "Dissociating Judgment from Response Processes in Statement Verification: The Effects of Experience on Each Component," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25 (January), 96–115.

Wason, P. C. (1960), "On the Failure to Eliminate Hypotheses in a Conceptual Task," *Quarterly Journal of Experimental Psychology*, 12 (August), 129–140.

Weaver, Charles A. III (1990), "Constraining Factors in the Calibration of Comprehension," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16 (March), 214–222.

——— (1993), "Do You Need a 'Flash' to Form a Flashbulb Memory?" *Journal of Experimental Psychology: General*, 122 (March), 39–46.

——— and Deborah S. Bryant (1995), "Monitoring of Comprehension: The Role of Text Difficulty in Metamemory for Narrative and Expository Text," *Memory & Cognition*, 23 (January), 12–22.

Weingardt, Kenneth R., R. Jacob Leonesio, and Elizabeth F. Loftus (1994), "Viewing Eyewitness Research from a Metacognitive Perspective," in *Metacognition: Knowing about Knowing*, ed. Janet Metcalfe and Arthur P. Shimamura, Cambridge, MA: MIT Press, 157–184.

———, Elizabeth F. Loftus, and D. Stephen Lindsay (1995), "Mis-

information Revisited: New evidence on the Suggestibility of Memory," *Memory & Cognition*, 23 (January), 72–82.

Weinstein, Neil D. (1980), "Unrealistic Optimism about Future Life Events," *Journal of Personality and Social Psychology*, 39 (November), 806–820.

Wells, Gary L. and Donna M. Murray (1984), "Eyewitness Confidence," in *Eyewitness Testimony: Psychological Perspectives*, ed. Gary L. Wells and Elizabeth F. Loftus, Cambridge: Cambridge University Press, 155–170.

West, Patricia M. (1996), "Predicting Preferences: An Examination of Agent Learning," *Journal of Consumer Research*, 23 (June), 68–80.

West, Richard F. and Keith E. Stanovich (1997), "The Domain Specificity and Generality of Overconfidence: Individual Differences in Performance Estimation Bias," *Psychonomic Bulletin and Review*, 4 (September), 387–392.

Wilson, Timothy D. and Nancy Brekke (1994), "Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations," *Psychological Bulletin*, 116 (July), 117–142.

———, Dana S. Dunn, Dolores Kraft, and Douglas J. Lisle (1989), "Introspection, Attitude Change, and Attitude-Behavior Consistency: The Disruptive Effects of Explaining Why We Feel the Way We Do," in *Advances in Experimental Social Psychology*, Vol. 22, ed. Leonard Berkowitz, San Diego, CA: Academic Press, 287–343.

——— and Suzanne J. LaFleur (1995), "Knowing What You'll Do: Effects of Analyzing Reasons on Self-Prediction," *Journal of Personality and Social Psychology*, 68 (January), 21–35.

———, Douglas J. Lisle, Jonathan W. Schooler, Sara D. Hodges, Kristen J. Klaaren, and Suzanne J. LaFleur (1993), "Introspecting about Reasons Can Reduce Post-Choice Satisfaction," *Personality and Social Psychology Bulletin*, 19 (June), 331–339.

——— and Jonathan W. Schooler (1991), "Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions," *Journal of Personality and Social Psychology*, 60 (February), 181–192.

Winman, Anders, Peter Juslin, and Mats Bjorkman (1998), "The Confidence-Hindsight Mirror Effect in Judgment: An Accuracy-Assessment Model for the Knew-It-All-Along Effect," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (March), 415–431.

Wood, Robert, Albert Bandura, and Trevor Bailey (1990), "Mechanisms Governing Organizational Performance in Complex Decision-Making Environments," *Organizational Behavior and Human Decision Processes*, 46 (August), 181–201.

Wright, George and Peter Ayton (1989), "Judgmental Probability Forecasts for Personal and Impersonal Events," *International Journal of Forecasting*, 5 (1), 117–125.

——— and Peter Ayton (1992), "Judgmental Probability Forecasting in the Immediate and Medium Term," *Organizational Behavior and Human Decision Processes*, 51 (April), 344–363.

———, Gene Rowe, Fergus Bolger, and John Gammack (1994), "Coherence, Calibration, and Expertise in Judgmental Probability Forecasting," *Organizational Behavior and Human Decision Processes*, 57 (January), 1–25.

Yates, J. Frank and Shawn P. Curley (1985), "Conditional Distribution Analyses of Probabilistic Forecasts," *Journal of Forecasting*, 4 (January–March), 61–73.

———, Ju-Whei Lee, and Julie G. Bush (1997), "General Knowledge Overconfidence: Cross-National Variations, Response Style, and 'Reality,'" *Organizational Behavior and Human Decision Processes*, 70 (May), 87–94.

———, Ju-Whei Lee, and Hiromi Shinotsuka (1996), "Beliefs about Overconfidence, Including Its Cross-National Variation," *Organizational Behavior and Human Decision Processes*, 65 (February), 87–94.

———, Ying Zhu, David L. Ronis, Deng-Feng Wang, Hiromi Shinotsuka, and Masanao Toda (1989), "Probability Judgment Accuracy: China, Japan, and the United States," *Organizational Behavior and Human Decision Processes*, 43 (April), 145–171.

Zajac, Edward J. and Max H. Bazerman (1991), "Blind Spots in Industry and Competitor Analysis: Implications of Interfirm (Mis)Perceptions for Strategic Decisions," *Academy of Management Review*, 16 (January), 37–56.

Zaragoza, Maria S. and Karen J. Mitchell (1996), "Repeated Exposure to Suggestion and the Creation of False Memories," *Psychological Science*, 7 (September), 294–300.

Zuckerman, Miron, Bella M. DePaulo, and Robert Rosenthal (1981), "Verbal and Nonverbal Communication of Deception," in *Advances in Experimental Social Psychology*, Vol. 14, ed. Leonard Berkowitz, New York: Academic Press, 1–35.

Zwick, Rami, Rik Pieters, and Hans Baumgartner (1995), "On the Practical Significance of Hindsight Bias: The Case of the Expectancy-Disconfirmation Model of Consumer Satisfaction," *Organizational Behavior and Human Decision Processes*, 64 (October), 103–117.