Overprecision in Judgment

Don A. Moore
University of California at Berkeley

Elizabeth R. Tenney
University of California at Berkeley

Uriel Haran
Ben-Gurion University of the Negev

Abstract: Overprecision in judgment is both the most durable and the least understood form of overconfidence. This chapter reviews the evidence on overprecision, highlighting its consequences in everyday life and for our understanding the psychology of uncertainty. There are some interesting explanations for overprecision, but none fully accounts for the diversity of the evidence. Overprecision remains an important phenomenon in search of a full explanation.

Moore, D. A., Tenney, E. R., & Haran, U. (in press). Overprecision in judgment. In G. Wu and G. Keren (Eds.), *Handbook of Judgment and Decision Making*. New York: Wiley

Key words: Overconfidence; overprecision; miscalibration

Address correspondence to don.moore@alumni.carleton.edu.

Overprecision in Judgment

Harold Camping foretold that the world would end on May 21st, 2011. At the time of his prophesy, Camping was the head of Family Radio, a Christian radio network. His reading of the Bible convinced him that the latter days were upon us and that the only way to be saved from the tribulation of Armageddon was to atone for one's sins and pledge oneself to God. Camping's followers dedicated themselves to warning others about the coming cataclysm and the urgent need to repent. Many gave up their jobs and donated all their savings—indeed all their worldly goods—to help fund the multi-million dollar publicity campaign to spread the word. When God disappointingly failed to bring the apocalypse, their faith was exposed as embarrassingly and tragically misplaced. How could they have been so sure of something so wrong?

The excessive faith that you know the truth is one form of overconfidence. And overconfidence may be the mother of all decision-making biases (Bazerman & Moore, 2013). We mean this in two ways. First, overconfidence is so durable and so ubiquitous. One popular text even claimed that, "No problem in judgment and decision making is more prevalent and more potentially catastrophic than overconfidence" (Plous, 1993, p. 217). Overconfidence has been blamed for, among many other things, the sinking of the Titanic, the nuclear accident at Chernobyl, the loss of Space Shuttles Challenger and Columbia, the subprime mortgage crisis of 2008 and the Great Recession that followed it, and the Deepwater Horizon oil spill in the Gulf of Mexico (Labib & Read, 2013; Moore & Swift, 2010). Daniel Kahneman (2011) called overconfidence "the most significant of the cognitive biases" and went so far as to call it the engine that drives capitalism. The second way in which overconfidence is the mother of all biases is that it gives other decision-making biases teeth. If we were appropriately humble about our vulnerabilities, then we might be able to better protect ourselves from the errors to which human judgment is prone. An excessive faith in the quality of our judgment—what we will call overprecision—leads us to rely on our own judgment too much, despite its many flaws.

The practical economic consequences of overprecision are plentiful and profound. Economic agents who are too sure they know what an asset is worth will not be as concerned as they should be about what the person on the other side of the trade knows, and consequently will be more willing to trade that asset than they should be (Daniel, Hirshleifer, & Sabrahmanyam, 2001). This overprecision in asset valuation is one potential explanation for the high rate of asset trading in stock markets and is the reason why overprecision may be the most important bias to affect economic markets (Odean, 1999; Rubinstein, 1985). People who are too sure they know what's going to happen will fail to protect themselves from risks whose probabilities they have underestimated (Silver, 2012). When people are too sure they know the truth they will neglect full consideration of others' perspectives and will fail to take advice that could have been helpful (Minson & Mueller, 2012; Yaniv, 2004). They may cling to beliefs that lack support, as did Harold Camping. This may manifest itself as egocentrism (M. Ross & Sicoly, 1979), the "false consensus" effect (Krueger & Clement, 1994; L. Ross, Greene, & House, 1977), or naïve realism (Pronin, Gilovich, & Ross, 2004; L. Ross & Ward, 1996). These tendencies lead people to believe that the way they see things is the only reasonable view and that others who disagree must be biased (Minson, Liberman, & Ross, 2009; Pronin, 2010).

Yet despite its well-documented effects and many published research studies, overprecision in judgment remains an effect lacking a full explanation. In this paper, we consider the evidence and the explanations, and we attempt to piece together the current state of our best understanding. But first, it is worth being specific about exactly what overprecision is and how we will focus this review.

Types of Overconfidence

Moore and Healy (2008) distinguish three varieties of overconfidence:

1) Overestimation is thinking that you're better than you are.
2) Overplacement is an exaggeration of the degree to which you are better than others.
3) Overprecision is the excessive faith that you know the truth.

This paper will focus on the third. We believe overprecision is the most interesting because it is both the least studied and the most robust form of overconfidence. The robustness claim rests on the fact that there are few, if any, documented reversals of overprecision, whereas there are numerous documented reversals of overestimation and overplacement (Erev, Wallsten, & Budescu, 1994; Moore, 2007). In other words, there are plentiful documented instances in which people believe that they are worse than they actually are (underestimation) or worse than others when they are not (underplacement) (Harris & Hahn, 2011; Moore, 2007). However, it is exceedingly rare for people to be less sure that they are right than they deserve to be. Obviously, overprecision does not require positivity or optimism. People are routinely too sure that bad things will happen, such as the world ending. To pick another tragic example, those who attempt suicide are usually too sure their lives will not improve.

What empirical evidence underlies our bold claim of universal overprecision? We review some of the evidence on overprecision in beliefs. This evidence comes from the lab and the field, from professionals and novices, with consequences ranging from the trivial to the tragic. The evidence reveals individuals' judgments to be overly precise—they are too sure they know the truth.

The 2AFC Paradigm

Laboratory studies of overprecision use several different paradigms for eliciting beliefs. The most common paradigm employs the 2-alternative forced choice approach, also known as 2AFC (Griffin & Brenner, 2004). Respondents see a question, choose between two possible answers, and indicate how confident they are that they have chosen correctly. For instance, the question might be, "Will the world end on May 21?" Respondents can answer yes or no and rate how confident they are that they are correct. Then, after May 21, we can compare confidence with actual outcomes and ask whether people were, on average, overconfident when their conviction led them to give away their worldly belongings before the anticipated Armageddon. The frequency with which a predicted outcome actually occurs is known as the "hit rate." Note that the comparison between confidence and hit rates becomes more useful as the number of questions increases. There is useful information present in a single report, but a large number of reports allows us to group together multiple reports and calibrate judgments by comparing average confidence across items (or items answered with a particular degree of confidence) to hit rate.

Calibration curves reveal the degree to which subjective confidence is matched by objective accuracy. Research finds again and again that subjective confidence is imperfectly correlated with accuracy (Keren, 1997; Lichtenstein & Fischhoff, 1980). When people are most confident, their confidence is not justified by accuracy. Perfect calibration implies a perfect correlation between confidence and accuracy. Instead, what often happens is the calibration curve is too flat: confidence is insufficiently sensitive to variation in accuracy.

Typically, researchers test for accuracy at given levels of confidence, but this approach biases the result in favor of results showing overconfidence: when confidence is high, confidence exceeds accuracy (Erev et al., 1994). The alternative approach would be to focus on hit rates, and ask how confident people are for given levels of accuracy. Analyzed this way, the imperfect correlation between confidence and accuracy produces what appears to be underconfidence: questions that people answer with high levels of accuracy are rarely reported with sufficient levels of confidence. It is easy to see why this is so at the extreme. For items that you get right, it is impossible to be overconfident (since confidence cannot exceed 100%) and you will appear underconfident. The more general point is that confidence and accuracy are often weakly correlated. Whether, on average, confidence exceeds accuracy for a given individual depends a great deal on how difficult the questions are (Juslin, Winman, & Olsson, 2000; Keren, 1988; Koriat, 2012). Hard questions tend to produce what appears to be overconfidence whereas easy questions produce underconfidence, both of which follow from regressiveness in confidence judgments.

The 2AFC paradigm has a number of notable limitations. One is that it confounds overprecision with overestimation of one's knowledge. It makes it impossible, for instance, to be too sure that one's estimate was incorrect. Another limitation is that this task does not capture continuous probability distributions very well. A third issue pits probabilistic against frequentistic reasoning in the data analysis. Participants must answer confidence questions on a scale that reflects degrees of probability for outcomes that will actually occur or not. There are some real philosophical questions about the comparability of these probability judgments with frequentistic counts (Gigerenzer, Hoffrage, & Kleinbölting, 1991; Macchi, 2000). Because we are concerned about these issues and because other useful reviews of the 2AFC literature already exist (Griffin & Brenner, 2004; Harvey, 1997; Hoffrage, 2004; Keren, 1991; Lichtenstein, Fischhoff, & Phillips, 1982), the 2AFC literature will not be our focus. Rather, we try to focus on studies that deliver insights into overprecision in judgment using other research paradigms. We believe that particularly useful lessons emerge from research that asks people to specify the precision of their beliefs more explicitly.

The Confidence-Interval Paradigm

An elicitation method that measures precision in judgment directly is the confidence interval. Alpert and Raiffa (1982) provided the methodological starting point for this literature when they asked their students to mark the particular fractiles in their subjective probability distributions. In particular, what they asked for were values associated with particular cumuluative probabilities. For instance, participants were asked to produce a low estimate of the number of eggs produced in the United States in 1965. That estimate should be low enough that there is only a 1% chance that the actual value fell below it. They were also asked for higher estimates—not just the 1st percentile, but the 25th, 50th, 75th, and 99th as well. They observed that their students described subjective probability distributions that covered too little of the event space—they set intervals that were simply too narrow and left too many possibilities out. Now, it's not just that these confidence intervals were too narrow in a statistically significant sense. They were whoppingly, catastrophically overprecise. 98% confidence intervals included the right answer, on average, only about 60% of the time.

Let's extrapolate from this remarkable result to imagine its implications. What if communities planned for floods, protecting themselves from 98% of the variation in high water levels, and this planning proved inadequate in 4 out of every 10 flood years? What if, when a corporation's strategic planners were 98% sure the company was ready to fulfill customer demand for a new product, it was wrong 40% of the time? What if, when people estimated that they were 98% sure they would finish a project by a certain date, they only actually completed their work punctually 60% of the time? What if political pollsters' 98% confidence intervals around a forecast election result were wrong 40% of the time? What if religious leaders who professed certainty about the coming apocalypse were wrong most of the time? Come to think of it, this might, in fact, be the world in which we live.

It's not just the 98% confidence intervals. Alpert and Raiffa also asked their subjects for the 25^{th} and 75^{th} fractiles (the distance between them accounting for 50% of the likely outcomes, and known fondly to statisticians everywhere as the interquartile range). These fractiles yielded intervals that were also too narrow, including the correct answer only a third of the time when they should have included the correct answer half the time.

Alpert and Raiffa were driven to exasperation with their respondents: "For heaven's sake, *Spread Those Extreme Fractiles*! Be honest with yourselves! Admit what you don't know!" (p. 301, emphasis in original), but this plea did not help much. Alpert and Raiffa's original demonstration has been replicated hundreds of times (Bazerman & Moore, 2013; Russo & Schoemaker, 1992). It remains one of the most reliable classroom demonstrations of decision biases. Ask your students for 90% confidence intervals around any ten numerical estimates, and you will get hit rates between 30% and 60%, suggesting they have drawn their confidence intervals too narrowly. They act as if they are surer than they deserve to be of their estimates.

Since Alpert and Raiffa (1982), a host of researchers have examined overprecision in judgment. This literature is marked by three particularly notable results, which we will review in more detail in the following sections. First is the robustness of the phenomenon. It is a satisfying effect to study because it appears so powerfully and consistently across populations, judgments, and contexts. It generalizes across cultures, genders, professions, ages, levels of expertise, and elicitation formats (Barber & Odean, 2001; Klayman, Soll, Gonzalez-Vallejo, & Barlas, 1999; McKenzie, Liersch, & Yaniv, 2008; Soll & Klayman, 2004). There are, however, some moderators of the size of the effect. These moderators provide some hints regarding the ultimate causes of overprecision in judgment and we review them below. At the same time, we note the shortage of moderators that can eliminate or reverse the effect—they can reduce it but they cannot make it go away.

The second striking feature of the results is that the width of 90% confidence intervals bears a weak relationship to other measures of confidence. For instance, investors' trading volume is not actually predicted very well by the overprecision of their 90% confidence intervals, and is better predicted by the belief that they are better than others (Glaser & Weber, 2007). This result is strange in part because it is at odds with economic models of the effects of overconfidence in financial markets (Daniel et al., 2001). We should be clear that this appears to be a shortcoming of this particular elicitation method—90% confidence intervals. The width of 90% confidence intervals also show low correlations with other behavioral measures of belief precision (Mannes & Moore, 2013). Indeed, the correlations between different measures of overprecision are so imperfect that they appear to measure different things.

This second issue leads us to the third and perhaps most important feature of the evidence: There is no single explanation that does a good job accounting for all of the research findings. There are various theories, and each can help account for some of the data, but the absence of a persuasive general theory suggests that overprecision in judgment is likely to be multiply determined.

Since Alpert and Raiffa's original demonstration, there have been innumerable powerful demonstrations of overprecision in judgment, from both the lab and the field. Next, we review some of this evidence by way of underscoring the durability and generality of the overprecision effect.

Ecological Evidence of Overprecision

There is no shortage of studies with students as participants, but it is also easy to find evidence of overprecision in professional judgment. A number of studies have examined physicians' tendency to gravitate toward a favored diagnosis and insufficiently consider other possibilities (Arkes, Wortmann, Saville, & Harkness, 1981; Christensen-Szalanski & Bushyhead, 1981; Hill, Gray, Carter, & Schulkin, 2005; Oskamp, 1965), but maybe selection and training encourage physician overconfidence. The same cannot be said for scientists, whose training emphasizes truth and the accurate assessment of degrees of uncertainty. Estimates of physical constants, such as the speed of light or Avogadro's number, are published with confidence intervals that reflect the researchers' uncertainty in the estimate. But even these confidence intervals contain the true value too rarely, suggesting the scientists are excessively confident in their estimates (Henrion & Fischhoff, 1986).

Investors, when they are too sure they know what an asset is worth, will be too eager to trade on that knowledge (Daniel, Hirshleifer, & Sabrahmanyam, 1998). It is, indeed, a mystery why investors trade so much, especially given the fact that trading is costly and timing the market is so difficult. Investor overconfidence may help account for this anomaly (Statman, Thorley, & Vorkink, 2006). The result is that those who trade more tend to perform worse (Barber & Odean, 2000). And when investors' errors correlate, overprecision can contribute to market volatility and speculative price bubbles (Scheinkman & Xiong, 2003). It would be nice if experienced market analysts could help investors avoid these biases, but they too appear vulnerable to overprecision (Stotz & von Nitzsch, 2005; Tyszka, Zielonka, & Dubra, 2002).

Inside of organizations, overprecision in judgment has a number of undesirable effects. One obvious one has to do with forecasting. Every consequential decision depends on a forecast. How many people to hire, how many factories to build, and how many widgets to produce all depend crucially on what future demand will be. There is ample evidence that organizational forecasts tend to be overly precise (Ben-David, Graham, & Harvey, 2013; Du, Budescu, Shelly, & Omer, 2011; Makridakis, Hogarth, & Gaba, 2009). If individuals are too sure of their forecasts, their planning will focus too tightly on a favored outcome and they will spend too little time planning for contingencies they believe to be unlikely (Aukutsionek & Belianin, 2001; Vallone, Griffin, Lin, & Ross, 1990). One consequence of such overprecision may be that people tend to search too little for ideas, people, and information (Bearden & Connolly, 2007; Haran, Ritov, & Mellers, 2013). They terminate their searches too early. And their overconfidence makes them less amenable to using normative decision aids like linear models that can improve their decisions (Sieck & Arkes, 2005).

We ought to be especially worried about leaders' proneness to overprecision. Overconfident candidates may be more likely to be selected for positions of leadership than ones who are more modest (Hayward, Shepherd, & Griffin, 2006). Any such selection effect is likely exacerbated by a treatment effect: the attention and adulation experienced by leaders may exacerbate overconfidence further (Malmendier & Tate, 2009; K. J. Murphy & Zabojnik, 2004). Managers who display overprecision do indeed introduce more risky products (Simon & Houghton, 2003) and put their firms at risk by undertaking too many risky projects and acquisitions (Malmendier & Tate, 2005, 2008).

Other paradigms

It is undeniable that specifying confidence intervals is unfamiliar to most people. One damning piece of evidence for this is that 98% confidence intervals are sometimes not much wider than 50% confidence intervals (Teigen & Jorgensen, 2005). The results suggest a form of attribute substitution (Kahneman & Frederick, 2002): Because it is too difficult to specify particular fractiles in their subjective probability distributions, people provide a rough range for a plausible "best guess." Nevertheless, having a good sense of our certainty is critically important in everyday life. We often have to make decisions that are affected by uncertainty. Let's say I have to get to an important lunch with my boss, but I am not sure how bad traffic will be. I leave on the early side so that I am more certain I will get there on time. Or let's say I want to avoid bouncing checks but I am not sure exactly how much is in my checking account. The solution is that I write fewer checks. When driving, if I am uncertain about how well my brakes are working, I can play it safe by staying well back of the car ahead of me.

These are all situations with asymmetric consequences. It is, for instance, worse to crash into the car ahead of you than to leave too much room between you. Asymmetric consequences drive us to be cautious by adjusting our response in the safe direction. The interesting implication is that it is possible to infer how certain someone is by how much they shift in response to the potential rewards or penalties. One study implemented this principle in an experiment that systematically varied the rewards for over- or underestimating the truth (Mannes & Moore, 2013). Research participants estimated temperatures in their city of residence on randomly-chosen dates from the past. In some rounds, they could maximize their payoffs by overestimating the actual temperature by no more than 8 degrees. In other rounds, they could maximize their payoffs by underestimating by the same amount. People responded to the different incentives by shifting their temperature estimates in the right direction, but this shifting was about half of what it should have been. Perceptual tasks show the same failure to shift in response to asymmetric payoffs (Mamassian, 2008).

These results provide important corroborative evidence by demonstrating overprecision without reliance on probabilities, confidence intervals, or other statistical concepts that research participants may not understand. Yes, people may misunderstand these things, but overprecision remains robust even in paradigms where their misunderstanding of statistical concepts cannot explain the result. It is possible that these misunderstandings contribute to the magnitude of overprecision in some findings, but even without it, the core finding persists: people behave as if they are too sure they know the truth.

Moderators of Overprecision

Despite the ubiquity of overprecision, research has identified some factors that moderate the severity of the bias. The way beliefs are elicited has big effects on how much confidence the judge will display in his or her answer. When you give people an interval and ask them to estimate how likely it is that the correct answer is inside it, they seem less confident than if you specify a probability of being right (say, 90%), and ask them for a confidence interval around it (Bolger & Harvey, 1995; Speirs-Bridge et al., 2010). In other words, the confidence people have in confidence intervals tends to be significantly lower than the confidence they have in probability estimates (Juslin, Wennerholm, & Olsson, 1999; Klayman, Soll, Juslin, & Winman, 2006). So, for instance, if you first ask people for 90% confidence intervals for ten items, and show those confidence intervals to others, they will tell you about 6 of them will include the correct answer (Cesarini, Sandewall, & Johannesson, 2006). Interestingly, this is not just because people think that others' answers are bad. If you ask people to estimate how many of their own 90% confidence intervals contain the correct answer, they similarly estimate that about 6 of them do (Cesarini et al., 2006).

Some have argued that this discrepancy in the confidence people convey is a consequence of the difference between probabilistic and frequentistic judgments (Gigerenzer & Hoffrage, 1995; Gigerenzer, 1993), but there are other crucial differences between assessing confidence at the level of an individual item vs. a set of items. Controlling for these confounds shrinks the difference between frequentistic and probabilistic modes dramatically (Griffin & Buehler, 1999).

It is reasonable to expect that knowledge and experience can have an effect on the degree of overconfidence people display. When a person is knowledgeable about a topic, she is more likely to make correct predictions about it and know the boundaries of her knowledge (Burson, Larrick, & Klayman, 2006; Kruger & Dunning, 1999). But with this expertise, and the feeling of knowing that accompanies it, confidence in her judgment increases as well (Mahajan, 1992). One study had expert and novice participants predict foreign exchange rates (Önkal, Yates, Simga-Mugan, & Öztin, 2003). Although experts' predictions were more accurate, experts also displayed higher confidence. Thus, experts and novices had similar levels of overconfidence. Other research corroborates this result (McKenzie et al., 2008): experts provided confidence intervals that were closer to the truth (measured by the distance between the intervals' midpoints and the true answers) but narrower than the intervals produced by novices. Higher accuracy on the one hand, and higher confidence on the other, cancels out effects of expertise on overprecision. Nevertheless, given the effectiveness of large amounts of practice accompanied by unambiguous, immediate feedback to reduce overconfidence in the 2AFC paradigm (Keren, 1987; A. H. Murphy & Winkler, 1977), we should not give up hope that some forms of expertise may help debias overprecision.

Another potential moderator may be the perceived importance of accuracy. There is some correlational evidence that estimates of items that are of personal importance to the judge (i.e., reasons that a judge finds influential for forming a decision; for example, reasons they want to vote for a given candidate) produce more overconfidence than those that lack personal importance (Paese & Feuer, 1991). It may be tempting to conclude that motivation contributes to the underlying cause of overprecision. However, what is conspicuously absent from this literature is any manipulation of the perceived importance of a topic with consequences on subsequent precision in judgment. Consistent with the shortage of evidence for wishful-thinking effects more generally (Krizan & Windschitl, 2007, 2009), and our own abject failures to find any evidence for motivational effects on overprecision in judgment (Logg, Haran, & Moore, 2013), we doubt that the motivation to believe in oneself is an important cause of overprecision. In one study, we gave participants questions from online intelligence tests. We told half the participants that they were just some math and logic puzzles. We told the other half it was an IQ test, and, just for good measure, we reminded them of a few of the positive life outcomes correlated with intelligence. The two groups showed no differences in their tendencies toward overprecision in estimating their own scores—or overestimation or overplacement, for that matter (Benoit, Dubra, & Moore, 2013; Logg et al., 2013).

Are some people more prone to overprecision than others? To date, research has found relatively little with regard to individual differences as predictors of overconfidence. There is some evidence that men are more overconfident in their estimates than women (Barber & Odean, 2001; Goodman-Delahunty, Granhag, Hartwig, & Loftus, 2010), although this difference is far from universal (Moore & Swift, 2010). Also extraverts tend to be more overconfident than introverts (Lynn, 1961; Schaefer, Williams, Goodie, & Campbell, 2004; Taylor & McFatter, 2003), but this effect was reversed in short-term recall tasks (Howarth & Eysenck, 1968; Osborne, 1972). Overconfidence has also been linked in some studies to proactiveness (Pallier et al., 2002), narcissism (Campbell, Goodie, & Foster, 2004), self-monitoring (Cutler & Wolfe, 1989), and trait optimism (Buehler & Griffin, 2003). However, these findings have yet to be followed up in a way that sheds much light on the durability of these correlations or provides much explanation for them.

Naïve intuition leads most people to expect cultural differences in overconfidence, such that Asians, being more self-deprecating (Akimoto & Sanbonmatsu, 1999), would be less likely to be overconfident. In fact, what evidence there is suggests the opposite: Asians are, if anything, more likely to display overprecision than are Europeans (Acker & Duck, 2008; Yates, Lee, & Bush, 1997; Yates, Lee, & Shinotsuka, 1996). In sum, the evidence on individual difference moderators of overprecision is complex, and more research is needed to determine how individual differences affect overconfidence (Moore & Healy, 2008; Moore & Swift, 2010).

Explanations

There are a number of theories that are worthy contenders as explanations for overprecision in judgment. Below, we review those we find most promising. As we will see, none of them is perfect.

Anchoring. Anchoring is often offered as a sensible and compelling explanation for overprecision. Tversky and Kahneman (1974) noted that confidence intervals are set too close to a "best estimate." The "best estimate" anchor becomes the starting point from which one fails to adjust sufficiently when estimating more extreme points in the probability distribution. Despite the plausibility of anchoring as an explanation for overprecision, it has weak empirical support. Several studies have shown that setting an anchor by first asking for a best guess (or a .50 fractile) not only fails to exacerbate overprecision, but sometimes actually accomplishes the opposite: *widening* confidence intervals (Block & Harper, 1991; Juslin et al., 1999; Selvidge, 1980; Soll & Klayman, 2004).

Conversational norms. One of the most compelling explanations for overprecision was offered by Yaniv and Foster (1995). They pointed out that the effort to be accurate comes at the expense of being informative. When your dinner companion informs you that she is 95% confident she will join you at the restaurant some time between 5:30 p.m. and 7:22 p.m., she is being accurate at the expense of being informative or useful for planning purposes. It may be more useful (and it is certainly more normal) for her to tell you precisely when she intends to get there, even if there is a chance she will be late.

A preference for informativeness over accuracy will have predictable effects on what types of communications are most credible and persuasive. Advisors, consultants, and would-be leaders will win our confidence by being precisely informative. Those who instead prefer to be accurate, admitting their uncertainty and widening the confidence intervals around their estimates for profitability, economic growth, or job creation, can find themselves sidelined. In Kahneman's (2011) words, "Experts who acknowledge the full extent of their ignorance may expect to be replaced by more confident competitors, who are better able to gain the trust of clients. An unbiased appreciation of uncertainty is a cornerstone of rationality—but it is not what people and organizations want" (p. 263)

People's preference for those who express confidence can indeed explain some of the dynamics of leader emergence (Anderson, Brion, Moore, & Kennedy, 2012; Van Swol & Sniezek, 2005). If people prefer confident leaders, then the leaders brave enough (or reckless enough) to express confidence in the hopes that no one will notice whether their confidence was warranted could very well win supporters and votes. And under some circumstances, would-be leaders who understand this dynamic may feel pressure to express more precision and greater confidence (Radzevick & Moore, 2011). Indeed, many of us may make the choice to be informative at the expense of being accurate. But it does not follow that most people express overprecision most of the time precisely because they hope to be more influential with others.

There are two main reasons to question whether conversational norms explain overprecision. First, it has difficulty accounting for why, in most studies where expressions of certainty amount to private communications with the experimenter, overprecision is so robust, especially when there is every reason for the subject to believe the experimenter is interested in accurate estimation. The conversational norms explanation for overprecision makes the clear prediction that the audience—and their goals—should matter. That is, when I am trying to decide when to book the dinner reservation, I should want my dinner companion to be informative. And when I am trying to decide how much steel reinforcement I need to support a new building, then I should want accuracy from the structural engineers. At least sometimes, researchers clearly ask their research participants for accuracy and instead get precision.

The second problem for the conversational norms explanation is that there is conspicuously little evidence in the literature that the audience (or its goals) affects how much overprecision people display. Researchers have documented similarly excessive precision, regardless of whether people are offering a private communication to the experimenter or a public commitment to the world. This absence of evidence dates back to Yaniv and Foster's (1995, 1997) original work, which showed that people routinely express a preference to receive informative over accurate communications, but not that communicators craft their claims of confidence with this preference in mind. Moreover, our own attempts to manipulate the audience have met with complete and utter failure (Haran, Radzevick, & Moore, 2010). The most extreme manipulations we could think of produced not a scintilla of statistically significant evidence for the notion that concern for the audience's interests or expectations affects expressions of overprecision.

Naïve intuitive statistician. Peter Juslin and his colleagues have offered an explanation they call the "naïve intuitive statistician" theory (Juslin, Winman, & Hansson, 2007). This theory is grounded in the statistical fact that small samples have smaller variances than the populations from which they are drawn. Human working memory is constrained and we cannot hold all relevant facts or estimates in our minds at one time. Because we just hold a small sample, we wind up underestimating the variance in the population. To put it another way, we underestimate the uncertainty in our knowledge because our attempts to estimate something produce less variance in estimates than they ought to. This is a clever theory, and it accounts for some evidence as Juslin et al. (2007) show, but not for the breadth and variety of evidence of overprecision, such as the persistence of overprecision when the number of facts or instances does not exceed human working memory (Mamassian, 2008; Moore & Healy, 2008).

Bias balance. Some have proposed that overconfidence may be a compensatory mechanism that offsets other biases. For instance, being overly sure we are right can help us overcome our tendency to be impatient and biased toward immediate gratification (Frederick, Loewenstein, & O'Donoghue, 2002). The logic of this argument goes as follows: Our big brains enable us to solve complicated problems, but sometimes these problems may require a great deal of patience. What prompts us to hang in there, toiling away on the big problems rather than going out and partying? The satisfaction that comes with knowing we're right—that we have solved the problem (Burton, 2008). The pleasure of being right, in this view, lures us to indulge in this self-satisfaction. Others have argued that overprecision in beliefs may help people overcome impatience or risk aversion (Kahneman & Lovallo, 1993). But these are not parsimonious explanations for the existence of overprecision. It is less efficient to design an organism with a bias and another bias to counteract it, than just to design one without either bias.

Related to this argument is the assertion that overprecision in judgment may, under some circumstances, confer adaptive benefits. One such theory holds that overconfidence can be useful for scaring off potential rivals and competitors (Charness, Rustichini, & van de Ven, 2011; Johnson & Fowler, 2011). However, this theory, as articulated, is more relevant to overplacement (the belief that one is better than others) than it is to overprecision. It is therefore not particularly helpful at explaining the overprecision that is our focus here. Another clever theory holds that overconfidence may give us courage and increase our self-efficacy. It may simply make us feel good or it may motivate action that benefits us (Benabou & Tirole, 2002). However, this theory is most relevant to overestimation (thinking that you are better than you are) and so is not all that helpful to our present purpose either.

There might also be a collective benefit from overconfidence. There can be little doubt, for instance, that the United States has benefited a great deal from the economic dynamism produced by energetic entrepreneurial activity. Great new companies have been created by those willing to throw caution to the wind and bet everything on a risky new venture. The economy may benefit from these individuals' risk-taking (Berg & Lein, 2005; Bernardo & Welch, 2001), and it may be the case that those who succeed are, on average, overconfident. But it does not follow that the individual is better off being more overconfident and gambling their savings starting a new business with a high probability of failure. Again, the explanation falls short of providing a persuasive account for the ubiquity of overprecision in judgment.

Existential know-it-all. Journalist Kathryn Schulz, in her provocative book, *Being Wrong* (2010), proposes one of the most beguiling explanations for overprecision in judgment that we have encountered. The explanation begins with the simple fact that people try to believe things they think are true. When they learn that something they believe is actually false, in that moment, they cease to believe it. It is in this sense, then, that people get used to being right about everything all the time: When reflecting on whether a particular belief is true, people, quite sensibly, conclude that, like everything they believe, it is true.

This explanation comes up short on scientific testability. The one opening seems to be the gray zone of approximation. There are lots of things we believe but know to be less than perfectly accurate. Most people would be willing to admit, for instance, that they do not know the exact length of the Nile River. We may know facts that are useful for estimating it, but we do not presume to have the number at the ready in our heads. Does our excessive faith in the relevance of our facts drive our overly precise estimates? Perhaps, and this may be why people are so much more confident about what they profess to know than about what others profess to know (Minson & Mueller, 2012; Tenney, 2013). But this issue deserves more research.

Misunderstanding. Avoiding overprecision requires an understanding, at some level, of probability distributions. Asking lay people to provide a probability distribution produces more confused looks than useful responses. Any simple attempt to explain to people what a probability distribution is will highlight the fact they do not usually think about uncertainty in terms of probability distributions, or if they do, they are amazingly good at hiding that understanding. We suspect that part of the answer to the durability of overprecision is that people make some consistent mistakes thinking about uncertainty that arise from their failure to understand probability distributions. For instance, they fail to center their confidence intervals on the most likely values (Moore, Carter, & Yang, 2013). In an error analogous to probability-matching, they center their confidence intervals on the wrong values but then overestimate the chances that they could be right.

Underprecision

Given the ubiquity of overprecision in judgment, we ought to be particularly interested in any findings of underprecision. While the literature does include a few studies that have reported instances of underprecision, or that have interpreted participants' behaviors as manifestations of underprecision, these rare outliers fall short of providing a useful explanation for the ubiquity of overprecision. For instance, some have claimed that auditors' professional judgments of a fictional auditing case do not show overprecision (Tomassini, Solomon, Romney, & Krogstad, 1982). These results, however, seem to be context-dependent, and have not generalized beyond the particular case study used in this particular study. Indeed, even auditors' judgments do appear overly precise on more standard tests of overprecision (Solomon, Ariyo, & Tomassini, 1985).

In another attempt to find underprecision, participants had a choice between betting on their own answers and betting on a gamble with the same probability of winning as the percentage of their correct answers (Blavatskyy, 2009). Participants showed a preference for the gamble of known probability. Interpreting this behavior as underprecision is certainly valid. However, the task involved a choice between two gambles with different levels of ambiguity, a factor which is known to affect people's choices and is just as plausible an explanation as underconfidence (see Heath & Tversky, 1990).

Several studies (Griffin, Tversky, Fischhoff, & Hall, 1992; Lichtenstein et al., 1982; Moore & Healy, 2008) explain underprecision with the notion of regressive beliefs—beliefs that are imperfectly correlated with the truth. They note that very easy questions, or questions for which there is a great deal of high-quality, relevant information, should be answered with very high levels of confidence. However, for these types of questions, the actual confidence subjects display is sometimes not high enough. The result is what appears to be underprecision.

Evidence for underprecision in confidence intervals has also been mixed. Participants in one study (Budescu & Du, 2007) showed underprecision when they provided 50% confidence intervals that were too wide, containing the correct answer 59% of the time. However, other researchers (Teigen & Jorgensen, 2005) found persistent overprecision even for 50% CIs (although it was less than that for 90% CIs). They found that both 90% and 50% CIs produced hit rates of about 23%. In line with this latter finding, researchers who manipulated the amount of evidence available for judges and found overprecision even among those who were aware that they knew next to nothing about the quantities they were asked to estimate. In other words, even those who knew they did not know the true answer did not stretch their confidence intervals wide enough to be calibrated with their confidence (Haran et al., 2013).

Because the width of confidence intervals is relatively insensitive to the confidence level, the obvious implication is that lowering the confidence level should produce underprecision. It does not really make sense to ask people for 20% confidence intervals because it is hard for them to know how to respond if they know the answer.

But there is a useful alternative: We can still ask for the 40th and 60th fractiles, which together define a range that ought to include the correct answer with 20% probability. Bob Clemen (personal communication, June 24, 2011) reports that asking people for 40th and 60th fractiles produces ranges that are too wide, in the sense that they contain the right answer more than 20% of the time. In other words, they set their 40th and 60th fractiles too far apart. This result, however, must be taken in conjunction with prior research showing that people set their 10th and 90th fractiles too close together (Soll & Klayman, 2004). Thus, if the 10th fractile is set too close to the middle, but the 40th fractile is set too low, then the interval between the 10th and 40th fracile is set too precisely. The same goes for the interval between the 60th and the 90th fractile. Again, what appears to be underprecision in setting small confidence intervals is likely to be a product of attribute substitution: People use more or less the same sensible lower-bound response when asked for either the 10th fractile and the 40th fractile, but whether this results in over- or under-precision depends on the standard set for them rather than their behavior.

Debiasing overprecision

Researchers have devoted a great deal of effort to developing ways to reduce overprecision. Most of the research has revolved around three main approaches:

1) Encouraging the consideration of more information and possible alternatives.
2) Decomposing the response set or alternatives into smaller components and considering each one of them separately.
3) Providing feedback.

The first approach to reducing overprecision is encouraging people to consider more information. It can help counteract the natural tendency, when attempting to answer a question, to focus on the first answer that comes to mind and ignore alternative outcomes (McKenzie, 1998). Getting the judge to consciously consider more information might then reduce overconfidence. Koriat, Lichtenstein, and Fischhoff (1980) had participants make estimates in a 2-alternative forced choice format, but before eliciting their confidence level in the accuracy of each choice, they asked participants to list arguments that contradicted their choices. This intervention successfully reduced overconfidence. Similarly, it appears to work to have people consider the alternative outcome (McKenzie, 1997) in the 2AFC paradigm, or to consider multiple alternatives before estimating the probability of an outcome (Hirt & Markman, 1995). However, this evidence comes from the 2AFC paradigm and its efficacy with confidence intervals or other measures of precision in judgment remains untested. Any such test ought to be cognizant of the evidence suggesting that more information can actually increase overconfidence under some circumstances. Providing information that is not diagnostic but that helps weave a coherent story can easily increase confidence more quickly than it increases accuracy (Koehler, Brenner, Liberman, & Tversky, 1996; Peterson & Pitz, 1988).

The second approach capitalizes on support theory's subaddativity effect (Tversky & Koehler, 1994). It suggests counteracting overprecision by taking the focal outcome and decomposing it into more specific alternatives. Fischhoff, Slovic, and Lichtenstein (1978) found that the sum of all probabilities assigned to the alternatives that make up the set is larger than the probability assigned to the set as a whole. Thus when estimating likelihoods for a number of possible outcomes, the more categories the judge is assessing (and the less we include under "all others"), the less confident they will be that their chosen outcome is the correct one.

Decomposition of confidence intervals has also achieved encouraging results. Soll and Klayman (2004) asked participants to estimate either an 80% confidence interval or the 10th and 90th fractiles separately (the distance between which should cover 80% of the participant's probability distribution). They found that the consideration of the high and low values separately resulted in wider and less overprecise intervals.

One elicitation method combines both the consideration of more information and the decomposition of the problem set into more specific subsets. The SPIES method (short for Subjective Probability Interval Estimates) (Haran, Moore, & Morewedge, 2010) turns a confidence interval into a series of probability estimates for different categories across the entire problem set. Instead of forecasting an interval that should include, with a certain level of confidence, the correct answer, the subject is presented with the entire range of possible outcomes. This range is divided into bins, and the subject estimates the probability of each bin to include the correct answer. For example, to predict the daily high temperature in Chicago on May 21st, we can estimate the probability that this temperature will be below 50ºF, between 51ºF and 60ºF, between 61ºF and 70ºF, between 71ºF and 80ºF, between 81ºF and 90ºF and 91ºF or more. Since these bins cover all possible options, the sum of all estimates should amount to 100%. From these subjective probabilities we can extract an interval for any desired confidence level. This method not only produces confidence intervals that are less overprecise than those produced directly, but it also reduces overprecision in subsequent estimates when subjects switch back to the traditional confidence interval method (Haran, Moore, et al., 2010). This reduction, however, does not seem to stem from the generalization of a better estimation process. Rather, the most pronounced improvements in estimates after a SPIES practice period seem to be when the SPIES task turns judges' attention to values previously regarded as the most unlikely (Haran, 2011). It may be possible, then, that when people are made aware of the possibility that their knowledge is incomplete (by directly estimating likelihoods of values which they completely ignored before), they increase caution in their confidence intervals.

The third approach to reducing overprecision is to provide the judge more feedback. It is tempting to believe that, given the size and consistency of the overprecision errors people make, providing them with feedback should provide obvious lessons that would allow them to correct their errors (González-Vallejo & Bonham, 2007; Krawczyk, 2011). However, the data are surprisingly mixed on this simple question. Feedback can indeed reduce overconfidence in some situations (Rakow, Harvey, & Finer, 2003; Subbotin, 1996). But the effect is far from universal, and the implication is that there are important moderators that determine whether feedback helps reduce overprecision biases (Benson & Onkal, 1992). For example, feedback results in improvement in low-probability judgments and more difficult tasks, but does not help, and sometimes even hurt performance of high-probability judgments and easy tasks (Baranski & Petrusic, 1999). Other studies found that the positive effect of feedback is much more pronounced after the first estimate a person makes, but its positive effect on calibration is gone thereafter (Baranski & Petrusic, 1994; Lichtenstein & Fischhoff, 1980). Benson and Onkal (1992) tested different kinds of feedback, and found that specific feedback about a forecaster's calibration can help, but simple outcome feedback does not. One of the key questions that all studies of overprecision must struggle with surrounds motivating people toward accuracy by rewarding it, and so now we turn our attention to scoring rules and incentives.

Incentive-compatible scoring rules for eliciting precision in judgments

One of the distinct challenges associated with the measure of overprecision in judgment is a shortage of incentive-compatible scoring rules. We would never claim that monetary compensation magically motivates rationality and accuracy in responding, but incentives can clarify instructions, at least in the following sense: they clarify what you (as a research participant) are supposed to do by specifying what you get rewarded for (Camerer & Hogarth, 1999; Hertwig & Ortmann, 2001). If you are being paid more to answer accurately, then you can show you're smart by responding accurately. Other motivations, such as responding according to conversational norms or social conventions, making yourself feel good, or trying to help the experimenter, will always be present to some degree. However, their relative importance can be reduced by increasing the payoff associated with accurate responding.

The Brier (1950) scoring rule (and its quadratic-scoring-rule alter-ego: Selten, 1998) are incentive compatible (so long as respondents are risk-neutral), but only work with probabilistic estimates of categorical outcomes. Researchers have employed them most often for estimates of binary outcome judgments (Hoffrage, 2004). While the study of confidence regarding beliefs about binary outcomes has produced a great many powerful and useful insights, they are necessarily constrained by the perfect confounding of overestimation and overprecision (Moore & Healy, 2008). These probabilistic estimates become substantially more informative when you break the state space up into mutually-exclusive categories and elicit subjective probabilities that the outcome will fall into each of the possible categories, as the SPIES elicitation method does (Haran, Moore, et al., 2010).

But the SPIES method forces the researcher to specify the intervals for a continuous state space. The arbitrary selection of category boundaries is likely to affect how individuals respond.

As Fox and Clemen (2005) show, survey respondents infer information from how the inquisitor divides up the state space (see also Sonnemann, Camerer, Fox, & Langer, 2013). In particular, respondents are reluctant to assign zero probability to a category the inquisitor has deemed important enough to ask about (Clemen & Ulu, 2008). This tendency does not always bias SPIES responses, but it is impossible to eliminate it as a concern.

The classic elicitation method for assessing precision in judgment—the 90% confidence interval—is not incentive-compatible. If you reward respondents for high hit-rates (getting the right answers inside their intervals) then clever respondents will make their intervals infinitely wide. If you reward respondents for providing narrower intervals, clever respondents will make their intervals infinitely narrow. If you try to reward both it becomes difficult to calibrate exactly how big this reward should be in order to perfectly counterbalance the incentive to increase hit rate, and the answer may depend on each respondent's subjective probability distribution.

Jose and Winkler (2009) saved the day by proposing an incentive-compatible scoring rule for continuous judgments. Their method only requires the inquisitor to specify a fractile, such as 10%. Thus armed, it becomes possible to ask people to estimate the fractile in a subjective probability distribution. For instance, if one were interested, one could ask respondents to estimate the 10th fractile of their subjective probability distribution of Barack Obama's body weight. Or one could ask the product manager responsible for the launch of the iPhone 6 for the 80th fractile of his subjective probability distribution of iPhone 6 sales in the first year following its launch. Of course, to get a full picture of the respondent's subjective probability distribution, it is useful to ask about a number of fractiles.

One obvious down-side of Jose and Winkler's approach requires the use of a payoff formula complex enough that few research participants will be able to look at and understand. And here we come to the annoying conundrum created by using incentive-compatible scoring rules. They are useful because they clarify what the respondent is supposed to do. But when they rely on complex mathematical formulae beyond the ken of most respondents, it undercuts the notion that their primary effect is to clarify. Instead, researchers find themselves prevailing on respondents to trust them. Moore and Healy (2008, p. 508) told their participants, "This formula may appear complicated, but what it means for you is very simple: You get paid the most when you honestly report your best guesses about the likelihood of each of the different possible outcomes." That's fine if respondents believed them. But those who suspected there may have been more going on were left with a more complex set of motivations.

Finally, we must note that the use of incentives is not any sort of panacea for getting research participants to pay attention and do what you, the researcher, want them to do. While monetary incentives can, under some circumstances, be helpful, and while they may increase your credibility with an economics audience, their benefits can be achieved by other means, including clear instructions, comprehension checks, or feedback. And they introduce other problems, including the challenge of explaining mathematical formulae that may be difficult for research participants to understand (Read, 2005).

(Mis)Perceiving Expressions of Confidence

Given the risks of relying on overconfident forecasts of economic growth, lovers' fidelity, or the end of the world, it could be helpful if we could learn to identify any of the three types of overconfidence in others. Do overconfident people act arrogant and haughty? Or does overconfidence masquerade as competence? Evidence suggests the latter is more likely to be the case. In one study (Anderson et al., 2012) participants made weight estimates of individuals in photographs and privately rated their own competence at the task compared to others. They then worked in groups to come to a consensus estimate. Participants who were overconfident that their knowledge was better than others' displayed behavioral cues consistent with competence. Group members likely had no idea who among them was competent and who was overconfident, and they rated overconfident peers as being more competent than those who were actually competent! Rampant overconfidence can thus go undetected, although there may be ways around this problem (e.g., Bonner & Bolinger, 2013).

In addition to the shortage of behavioral cues to detect overconfidence, people are likely to be quick to assume that displays of confidence are meaningful and are warranted. According to the presumption of calibration hypothesis (Tenney, Spellman, & MacCoun, 2008; Tenney & Spellman, 2011), the default assumption is that the people around us are well-calibrated (i.e., not overconfident). That is, without evidence to the contrary, people assume others have insight into the quality of their own knowledge and that their displayed confidence is a good predictor of accuracy. Of course, certain baseline assumptions would override this default. Some people are naturally skeptical or will use top-down information or stereotypes (e.g., thinking all politicians are overconfident) when judging others. Barring assumptions like those, according to the hypothesis, the typical assumption is that others are well-calibrated and confidence signifies accuracy.

But what happens when people have new information that could contradict the initial assumption? Some researchers have posited a simple confidence heuristic that leads people to trust whoever is most confident (Price & Stone, 2004; Thomas & McFadyen, 1995). However, the presumption of calibration hypothesis takes a different view. It predicts that if people have clear information about the link between confidence and accuracy for a given person (i.e., calibration), then they will use it. For example, if people recognize that someone is overconfident, then confident statements from that person will lose credibility, which could affect whether the person is hired or believed in court (Spellman & Tenney, 2010; Tenney, MacCoun, Spellman, & Hastie, 2007; Tenney et al., 2008). Believing that someone is overconfident will undermine their credibility, as with the boy who cried wolf. Nevertheless, there may be times when overconfidence might be desirable—such as when people need motivation to succeed (Tenney, Logg, & Moore, 2013).

Note that there is an important distinction between people's general demeanor (e.g., extraverted, dominant, likeable) and people's confidence in specific claims (which can be warranted or not). If someone acts assertive (in general) and overconfident (about some particular fact), people might give this person social credit for being assertive but nevertheless trust his or her confident claims less than had this person been both assertive and well-calibrated. Examining confidence expressions in naturalistic settings make it difficult to distinguish these two. The negative effects of overly precise claims could be masked by or confounded with positive effects of a confident demeanor (e.g., Anderson et al., 2012; Kennedy, Anderson, & Moore, 2013). Future research should seek to clarify the causal role of these two different forms of confidence expression.

According to the presumption of calibration hypothesis, people do not trust confidence as much when they have reason to believe that confidence is uninformative or excessive. This research suggests that generally people do not *like* or *want* overconfidence when they can spot it in others. An important caveat is that sometimes people do not spot it. Making inferences about another's calibration could be a sophisticated process that people are not always able to engage. Adults under cognitive load used information about calibration less than adults not under cognitive load, and children (ages 5 and 6) did not use calibration (Tenney, Small, Kondrad, Jaswal, & Spellman, 2011). Similarly, adults who had to pay to acquire calibration information often did not do so (Sah, Moore, & MacCoun, 2013), suggesting that when calibration information is difficult to obtain or difficult to mentally process, the default assumption remains intact. Thus, it can often look as though people are using a confidence heuristic.

Earlier, we quoted Kahneman's (2011) assertion that people give more credibility to confident experts. He wrote that "An unbiased appreciation of uncertainty is a cornerstone of rationality—but it is not what people and organizations want" (p. 263). Here, we must add an addendum to this claim. People and organizations *do* want unbiased experts who appreciate uncertainty, but people are nevertheless drawn to the confident experts, especially when they believe that high confidence is a signal of high performance (e.g., Tenney & Spellman, 2011). Elsewhere, Kahneman offered the following advice: "You should not take assertive and confident people at their own evaluation unless you have independent reason to believe that they know what they are talking about. Unfortunately, this advice is difficult to follow: overconfident professionals sincerely believe they have expertise, act as experts and look like experts. You will have to struggle to remind yourself that they may be in the grip of an illusion" (Kahneman, 2011a). That is the illusion of overprecision—the unjustified feeling of knowing.

Future research

This review should have made clear that overprecision is one of the largest, most durable, and most important effects in the literatures on human judgment, heuristics and biases, and behavioral decision research. Yet it is an effect that remains in need of a full explanation. This is a vexing situation that we hope will stimulate research into the ultimate causes of overprecision in judgment. The holy grail of any such research program would be a single explanation that could account persuasively for the breadth and variety of overprecision results documented in the literature. However, we may have to content ourselves with the conclusion that no one explanation can account for all the evidence of overprecision. It seems likely that the phenomenon is multiply determined, and that there are several explanations for it, each one accounting for extant evidence to varying degrees.

In the past, the study of overprecision has relied too heavily on a small number of paradigms that bear little relationship to common everyday judgments. We see opportunities to study overprecision in naturalistic contexts where it affects behavior. For instance, focusing on reducing overprecision bias could have important benefits for conflict resolution. Overprecision gets in the way of appreciating others' perspectives and resolving disagreements (Loewenstein & Moore, 2004; Minson et al., 2009; Thompson & Loewenstein, 1992). Indeed, there is ample evidence of this problem: There are too many legal disputes that go on too long (Priest & Klein, 1984), too many labor strikes (Babcock, Wang, & Loewenstein, 1996), and too many violent conflicts (Johnson, 2004) because people are too sure that their view of reality is the correct one, or that the outcome they believe will happen will indeed happen. There is preliminary evidence that people recognize overprecision in others (and discredit others' later confident assertions as a result) but fail to do the same for themselves when given the same information about their own overprecision (Tenney, 2013). Experiments and interventions that work to help people to recognize when they are overprecise could add valuable insight into the causes and consequences of overprecision—and could potentially help opposing parties in a variety of situations see eye to eye.

Another fertile and important context in which to study overprecision is forecasting. Every consequential decision depends, to some degree, on a forecast. Corporations forecast product demand to help them plan for hiring, investments in productive capacity, and production quantities. The nation's intelligence agencies forecast world events, the fall of foreign regimes, and the prices of global commodities in order to help develop the most effective foreign policies (Tetlock & Mellers, 2011). Individuals forecast their future incomes, expenditures, and consumption in order to plan family budgets. Forecasting well sets the stage to maximize wise decision making and future welfare. Forecasting questions seem particularly amenable to the use of fractile elicitations, and hold tremendous promise, both to produce useful research insights into the nature of overprecision, and also to help people and organizations plan more effectively.

Coda

Harold Camping's message was, in a twisted sense, a spectacularly successful act of leadership. Few among our business or political leaders can inspire the kind of devotion he elicited from his followers. Camping's success arose in part from one key element of successful leaders: confidence. Confidence makes it look like leaders know what they're doing. They plot a path for others to follow and articulate the glorious rewards that will come to those who stay the course.

Those who are the most certain of themselves are most likely to attain positions of status in their groups (Anderson et al., 2012). Group status hierarchies, once established, are fairly resistant to change (Kilduff & Galinsky, 2013; Savin-Williams, 1979). Indeed, although he faced extreme ridicule and lost credibility in dramatic fashion, Harold Camping remained at the head of his Family Radio network even after his 2011 prophesy proved false.

Although we may be tempted to shake our heads in pity at those who blindly follow leaders who express excessive certainty, we would do well to offer ourselves the same sympathies. After all, most of us find confidence in others to be compelling and inspiring, especially when we are too busy to think about whether their confidence is warranted. The alternative, which all of us would do well to consider, is to select leaders who are accurate. Yes, confidence is nice, but better still is to place our trust in those who actually know the truth and are willing to consider, and prepare for, many possibilities when they do not. And that requires that we grow more comfortable with the inherent uncertainty in predicting the future, especially when it comes to the apocalypse.

References

Acker, D., & Duck, N. W. (2008). Cross-cultural overconfidence and biased self-attribution. *Journal of Socio-Economics*, *37*(5), 1815–1824.

Akimoto, S. A., & Sanbonmatsu, D. M. (1999). Differences in Self-Effacing Behavior between European and Japanese Americans; Effect on Competence Evaluations. *Journal of Cross-Cultural Psychology*, *30*(2), 159–177.

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Anderson, C., Brion, S., Moore, D. A., & Kennedy, J. A. (2012). A social-functional account of overconfidence. *Journal of Personality and Social Psychology*, *103*(4), 718–735.

Arkes, H. R., Wortmann, R. L., Saville, P. D., & Harkness, A. R. (1981). Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology*, *66*(2), 252–254.

Aukutsionek, S. P., & Belianin, A. V. (2001). Quality of forecasts and business performance: A survey study of Russian managers. *Journal of Economic Psychology*, *22*(5), 661–692.

Babcock, L., Wang, X., & Loewenstein, G. (1996). Choosing the wrong pond: Social comparisons in negotiations that reflect a self-serving bias. *Quarterly Journal of Economics*, *111*(1), 1–19.

Baranski, J. V, & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception & Psychophysics*, *55*(4), 412–428 LA – English. doi:10.3758/BF03205299

Baranski, J. V, & Petrusic, W. M. (1999). Realism of confidence in sensory discrimination. *Perception & Psychophysics*, *61*(7), 1369–1383.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance*, *55*(2), 773–806.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261–292. doi:10.1162/003355301556400

Bazerman, M. H., & Moore, D. A. (2013). *Judgment in managerial decision making* (8th ed.). New York: Wiley.

Bearden, J. N., & Connolly, T. (2007). Multi-attribute sequential search. *Organizational Behavior & Human Decision Processes*, (1), 147–158.

Benabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly Journal of Economics*, *117*(3), 871–915.

Ben-David, I., Graham, J. R., & Harvey, C. R. (2013). Managerial miscalibration. *Quarterly Journal of Economics*, *In press*.

Benoit, J.-P., Dubra, J., & Moore, D. A. (2013). Does the Better-Than-Average Effect Show That People Are Overconfident?: Two Experiments. *Unpublished manuscript2*.

Benson, P. G., & Onkal, D. (1992). The effects of feedback and training on the performance of probability forecasters. *International Journal of Forecasting*, *8*(4), 559–573. doi:10.1016/0169-2070(92)90066-I

Berg, N., & Lein, D. (2005). Does society benefit from investor overconfidence in the ability of financial market experts? *Journal of Economic Behavior & Organization*, *58*(1), 95–116.

Bernardo, A. E., & Welch, I. (2001). On the Evolution of Overconfidence and Entrepreneurs. *Journal of Economics and Management Strategy*, *10*(3), 301–330.

Blavatskyy, P. R. (2009). Betting on own knowledge: Experimental test of overconfidence. *Journal of Risk and Uncertainty*, *38*(1), 39–49. doi:10.1007/s11166-008-9048-7

Block, R. A., & Harper, D. R. (1991). Overconfidence in estimation: Testing the anchoring-and-adjustment hypothesis. *Organizational Behavior and Human Decision Processes*, *49*(2), 188–207.

Bolger, F., & Harvey, N. (1995). Judging the probability that the next point in an observed time-series will be below, or above, a given value. *Journal of Forecasting*, *14*(7), 597–607.

Bonner, B. L., & Bolinger, A. R. (2013). Separating the confident from the correct: Leveraging member knowledge in groups to improve decision making and performance. *Organizational Behavior and Human Decision Processes*, *122*(2), 214–221.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *78*(1), 1–3.

Budescu, D. V, & Du, N. (2007). The coherence and consistency of investors' probability judgments. *Management Science*, *53*(11), 1731–1745.

Buehler, R., & Griffin, D. W. (2003). Planning, personality, and prediction: The role of future focus in optimistic time predictions. *Organizational Behavior and Human Decision Processes*, *92*(1-2), 80–90. doi:10.1016/S0749-5978(03)00089-X

Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, *90*(1), 60–77.

Burton, R. A. (2008). *On being certain: Believing you are right even when you're not*. New York: St. Martin's press.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, *19*(1-3), 7–42.

Campbell, W. K., Goodie, A. S., & Foster, J. D. (2004). Narcissism, confidence, and risk attitude. *Journal of Behavioral Decision Making*, *17*(4), 481–502.

Cesarini, D., Sandewall, Ö., & Johannesson, M. (2006). Confidence interval estimation tasks and the economics of overconfidence. *Journal of Economic Behavior and Organization*, *61*(3), 453–470.

Charness, G., Rustichini, A., & van de Ven, J. (2011). Overconfidence, self-esteem, and strategic deterrence. *Unpublished Manuscript*.

Christensen-Szalanski, J. J., & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 928–935.

Clemen, R. T., & Ulu, C. (2008). Interior additivity and subjective probability assessment of continuous variables. *Management Science*, *54*(4), 835–851.

Cutler, B. L., & Wolfe, R. N. (1989). Self-monitoring and the association between confidence and accuracy. *Journal of Research in Personality*, *23*(4), 410–420. doi:10.1016/0092-6566(89)90011-1

Daniel, K. D., Hirshleifer, D. A., & Sabrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *Journal of Finance*, *53*(6), 1839–1885.

Daniel, K. D., Hirshleifer, D. A., & Sabrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *Journal of Finance*, *56*(3), 921–965.

Du, N., Budescu, D. V, Shelly, M., & Omer, T. C. (2011). The appeal of vague financial forecasts. *Organizational Behavior & Human Decision Processes*, *114*(2), 179–189.

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review*, *101*(3), 519–527.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(2), 330–344.

Fox, C. R., & Clemen, R. T. (2005). Subjective probability assessment in decision analysis: Partition dependence and bias toward the ignorance prior. *Management Science*, *51*(9), 1417.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401.

Gigerenzer, G. (1993). The bounded rationality of probabilistic mental modules. In K. I. Manktelow & D. E. Over (Eds.), *Rationality* (pp. 127–171). London: Routledge.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*(4), 684–704. doi:10.1037/0033-295X.102.4.684

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*(4), 506–528.

Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *Geneva Risk and Insurance Review*, *32*, 1–36.

González-Vallejo, C., & Bonham, A. (2007). Aligning confidence with accuracy: revisiting the role of feedback. *Acta Psychologica*, *125*(2), 221–39. doi:10.1016/j.actpsy.2006.07.010

Goodman-Delahunty, J., Granhag, P. A., Hartwig, M., & Loftus, E. F. (2010). Insightful or wishful: Lawyers' Ability to Predict Case Outcomes. *Psychology, Public Policy, and Law*, *16*(2), 133–157.

Griffin, D. W., & Brenner, L. (2004). Perspectives on probability judgment calibration. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 177–199). Malden, MA: Blackwell.

Griffin, D. W., & Buehler, R. (1999). Frequency, probability, and prediction: Easy solutions to cognitive illusions? *Cognitive Psychology*, *38*(1), 48–78.

Griffin, D. W., Tversky, A., Fischhoff, B., & Hall, J. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*, 411–435.

Haran, U., Moore, D. A., & Morewedge, C. K. (2010). A simple remedy for overprecision in judgment. *Judgment and Decision Making*, *5*(7), 467–476.

Haran, U., Radzevick, J. R., & Moore, D. A. (2010). Audience effects on overprecision in judgment. *Unpublished Manuscript*.

Haran, U., Ritov, I., & Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Unpublished Manuscript*.

Harris, A. J. L., & Hahn, U. (2011). Unrealistic optimism about future life events: A cautionary note. *Psychological Review*, *118*(1), 135–154.

Harvey, N. (1997). Confidence in judgment. *Trends in Cognitive Sciences*, *1*(2), 78–82.

Hayward, M. L. A., Shepherd, D. A., & Griffin, D. W. (2006). A hubris theory of entrepreneurship. *Management Science*, *52*(2), 160–172.

Heath, C., & Tversky, A. (1990). Preference and belief: Ambiguity and competence in choice under uncertainty. In *Contemporary issues in decision making* (pp. 93–123). Amsterdam, Netherlands: North-Holland.

Henrion, M., & Fischhoff, B. (1986). Assessing uncertainty in physical constants. *American Journal of Physics*, *54*(9), 791–798.

Hertwig, R., & Ortmann, A. (2001). Experimental practies in economics: A methodological challege for psychologists? *Behavioral and Brain Sciences*, *24*, 383–451.

Hill, L. D., Gray, J. J., Carter, M. M., & Schulkin, J. (2005). Obstetrician-gynecologists' decision making about the diagnosis of major depressive disorder and premenstrual dysphoric disorder. *Journal of Psychosomatic Obstetrics & Gynecology*, *26*(1), 41–51.

Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, *69*(6), 1069.

Hoffrage, U. (2004). Overconfidence. In R. F. Pohl (Ed.), *Cognitive illusions: Fallacies and biases in thinking, judgment, and memory* (pp. 235–254). Hove, UK: Psychology Press.

Howarth, E., & Eysenck, H. J. (1968). Extraversion, arousal, and paired-associate recall. *Journal of Experimental Research in Personality*, *3*(2), 114–116.

Johnson, D. D. P. (2004). *Overconfidence and War: The Havoc and Glory of Positive Illusions*. Cambridge, MA: Harvard University Press.

Johnson, D. D. P., & Fowler, J. H. (2011). The evolution of overconfidence. *Nature*, *477*(7364), 317–320.

Jose, V. R. R., & Winkler, R. L. (2009). Evaluating quantile assessments. *Operations Research*, *57*(5), 1287–1297.

Juslin, P., Wennerholm, P., & Olsson, H. (1999). Format dependence in subjective probability calibration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 1038–1052.

Juslin, P., Winman, A., & Hansson, P. (2007). The naïve intuitive statistician: A naïve sampling model of intuitive confidence intervals. *Psychological Review*, *114*(3), 678–703.

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. *Psychological Review*, *107*(2), 384–396.

Kahneman, D. (2011a). Don't blink! The hazards of confidence. *New York Times*.

Kahneman, D. (2011b). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.

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

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk and risk taking. *Management Science*, *39*, 17–31.

Kennedy, J. A., Anderson, C., & Moore, D. A. (2013). When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence. *Organizational Behavior and Human Decision Processes*, *122*(2), 266–279.

Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes*, *39*(1), 98–114.

Keren, G. (1988). On the ability of monitoring non-veridical perceptions and uncertain knowledge: Some calibration studies. *Acta Psychologica*, *67*(2), 95–119.

Keren, G. (1991). Calibration and probability judgments: Conceptual and methodological issues. *Acta Psychologica*, *77*, 217–273.

Keren, G. (1997). On the calibration of probability judgments: Some critical comments and alternative perspectives. *Journal of Behavioral Decision Making*, *10*(3), 269–278.

Kilduff, G. J., & Galinsky, A. D. (2013). From the ephemeral to the enduring: How approach-oriented mindsets lead to greater status. *Journal of Personality and Social Psychology*, *105*(5), 816.

Klayman, J., Soll, J. B., Gonzalez-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, *79*(3), 216–247.

Klayman, J., Soll, J. B., Juslin, P., & Winman, A. (2006). Subjective confidence and the sampling of knowledge. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition*. New York: Cambridge University Press.

Koehler, D. J., Brenner, L. A., Liberman, V., & Tversky, A. (1996). Confidence and accuracy in trait inference: Judgment by similarity. *Acta Psychologica*, *92*(1), 33–57.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, *119*(1), 80–113.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 107–118.

Krawczyk, M. (2011). Overconfident for real? Proper scoring for confidence intervals. *Unpublished Manuscript*.

Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological Bulletin*, *133*(1), 95–121.

Krizan, Z., & Windschitl, P. D. (2009). Wishful Thinking about the Future: Does Desire Impact Optimism? *Social and Personality Psychology Compass*, *3*(3), 227–243. doi:10.1111/j.1751-9004.2009.00169.x

Krueger, J. I., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, *67*(4), 596–610.

Kruger, J., & Dunning, D. (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*(6), 1121–1134.

Labib, A., & Read, M. (2013). Not just rearranging the deckchairs on the Titanic: Learning from failures through risk and reliability analysis. *Safety Science*, *51*(1), 397–413.

Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational Behavior and Human Decision Processes*, *26*(2), 149–171.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art in 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–333). Cambridge, England: Cambridge University Press.

Loewenstein, G., & Moore, D. A. (2004). When ignorance is bliss: Information exchange and inefficiency in bargaining. *Journal of Legal Studies*, *33*(1), 37–58.

Logg, J. M., Haran, U., & Moore, D. A. (2013). Motivation and overconfidence. *Unpublished Manuscript*.

Lynn, R. (1961). Introversion-extraversion differences in judgments of time. *Journal of Abnormal and Social Psychology*, *63*(2), 457–8.

Macchi, L. (2000). Partitive Formulation of Information in Probabilistic Problems: Beyond Heuristics and Frequency Format Explanations. *Organizational Behavior and Human Decision Processes*, *82*(2), 217–236. doi:http://dx.doi.org/10.1006/obhd.2000.2895

Mahajan, J. (1992). The overconfidence effect in marketing management predictions. *Journal of Marketing Research*, *29*(3), 329–342.

Makridakis, S., Hogarth, R. M., & Gaba, A. (2009). Forecasting and uncertainty in the economic and business world. *International Journal of Forecasting*, *25*(4), 794–812.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance*, *60*(6), 2661–2700.

Malmendier, U., & Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, *89*(1), 20–43.

Malmendier, U., & Tate, G. (2009). Superstar CEOs. *The Quarterly Journal of Economics*, *124*(4), 1593–1638.

Mamassian, P. (2008). Overconfidence in an objective anticipatory motor task. *Psychological Science*, *19*(6), 601–606.

Mannes, A. E., & Moore, D. A. (2013). A behavioral demonstration of overconfidence in judgment. *Psychological Science*, *24*(7), 1190–1197.

McKenzie, C. R. M. (1997). Underweighting alternatives and overconfidence. *Organizational Behavior and Human Decision Processes*, *71*(2), 141–160. doi:10.1006/obhd.1997.2716

McKenzie, C. R. M. (1998). Taking into account the strength of an alternative hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 771–792. doi:10.1037/0278-7393.24.3.771

McKenzie, C. R. M., Liersch, M. J., & Yaniv, I. (2008). Overconfidence in interval estimates: What does expertise buy you? *Organizational Behavior and Human Decision Processes*, *107*, 179–191.

Minson, J. A., Liberman, V., & Ross, L. (2009). Two to tango: The effect of collaborative experience and disagreement on dyadic judgment. *Personality and Social Psychology Bulletin*, *37*, 1325–1338.

Minson, J. A., & Mueller, J. (2012). The cost of collaboration: Why joint decision making exacerbates rejection of outside information. *Psychological Science*, *23*(3), 219–224.

Moore, D. A. (2007). Not so above average after all: When people believe they are worse than average and its implications for theories of bias in social comparison. *Organizational Behavior and Human Decision Processes*, *102*(1), 42–58.

Moore, D. A., Carter, A., & Yang, H. H. J. (2013). Overprecision and the Quincunx. *Unpublished Manuscript*. Retrieved from http://learnmoore.org/mooredata/BDE/

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502–517.

Moore, D. A., & Swift, S. A. (2010). The Three Faces of Overconfidence in Organizations. In R. Van Dick & J. K. Murnighan (Eds.), *Social Psychology of Organizations* (pp. 147–184). Oxford: Taylor & Francis.

Murphy, A. H., & Winkler, R. L. (1977). Can weather forecasters formulate reliable probability forecasts of precipitation and temperature? *National Weather Digest*, *2*, 2–9.

Murphy, K. J., & Zabojnik, J. (2004). CEO pay and appointments: A market-based explanation for recent trends. *American Economic Review*, *94*(2), 192–196.

Odean, T. (1999). Do investors trade too much? *American Economic Review*, *89*(5), 1279–1298.

Önkal, D., Yates, J. F., Simga-Mugan, C., & Öztin, S. (2003). Professional vs. amateur judgment accuracy: The case of foreign exchange rates. *Organizational Behavior and Human Decision Processes*, *91*(2), 169–185.

Osborne, J. W. (1972). Short- and long-term memory as a function of individual differences in arousal. *Perceptual and Motor Skill*, *34*, 587–593.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, *29*(3), 261–265.

Paese, P. W., & Feuer, M. A. (1991). Decisions, actions, and the appropriateness of confidence in knowledge. *Journal of Behavioral Decision Making*, *4*(1), 1–16.

Pallier, G., Wilkinson, R., Danthir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *Journal of General Psychology*, *129*(3), 257–299.

Peterson, D. K., & Pitz, G. F. (1988). Confidence, uncertainty, and the use of information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 85.

Plous, S. (1993). *The Psychology of Judgment and Decision Making*. New York: McGraw-Hill.

Price, P. C., & Stone, E. R. (2004). Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic. *Journal of Behavioral Decision Making*, *17*, 39–57.

Priest, G. L., & Klein, B. (1984). The selection of disputes for litigation. *The Journal of Legal Studies*, *13*(1), 1–55.

Pronin, E. (2010). The introspection illusion. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 41, pp. 1–67). Burlingon: Academic Press.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. *Psychological Review*, *111*(3), 781–799.

Radzevick, J. R., & Moore, D. A. (2011). Competing to be certain (but wrong): Social pressure and overprecision in judgment. *Management Science*, *57*(1), 93–106.

Rakow, T., Harvey, N., & Finer, S. (2003). Improving calibration without training: the role of task information. *Applied Cognitive Psychology*, *17*(4), 419–441.

Read, D. (2005). Monetary incentives, what are they good for? *Journal of Economic Methodology*, *12*(2), 265–276.

Ross, L., Greene, D., & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301.

Ross, L., & Ward, A. (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In E. Reed, E. Turiel, & T. Brown (Eds.), *Values and knowledge* (pp. 103–135).

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, *37*, 322–336.

Rubinstein, A. (1985). A bargaining model with incomplete information about time preferences. *Econometrica*, *53*(5), 1151–1172.

Russo, J. E., & Schoemaker, P. J. H. (1992). Managing overconfidence. *Sloan Management Review*, *33*(2), 7–17.

Sah, S., Moore, D. A., & MacCoun, R. J. (2013). Cheap talk and credibility: The consequences of confidence and accuracy on advisor credibility and persuasiveness. *Organizational Behavior and Human Decision Processes*, *121*(2), 246–255.

Savin-Williams, R. C. (1979). Dominance hierarchies in groups of early adolescents. *Child Development*, *50*(4), 923–935.

Schaefer, P. S., Williams, C. C., Goodie, A. S., & Campbell, W. K. (2004). Overconfidence and the big five. *Journal of Research in Personality*, *38*(5), 473–480.

Scheinkman, J. A., & Xiong, W. (2003). Overconfidence and Speculative Bubbles. *Journal of Political Economy*, *111*(6), 1183–1219.

Schulz, K. (2010). *Being wrong*. New York: Ecco.

Selten, R. (1998). Axiomatic characterization of the quadratic scoring rule. *Experimental Economics*, *1*(1), 43–61.

Selvidge, J. E. (1980). Assessing the extremes of probability distributions by the fractile method. *Decision Sciences*, *11*(3), 493–502.

Sieck, W. R., & Arkes, H. R. (2005). The recalcitrance of overconfidence and its contribution to decision aid neglect. *Journal of Behavioral Decision Making*, *18*(1), 29–53.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail--but some don't*. Penguin Press.

Simon, M., & Houghton, S. M. (2003). The relationship between overconfidence and the introduction of risky products: Evidence from a field study. *Academy of Management Journal*, *46*(2), 139–150.

Soll, J. B., & Klayman, J. (2004). Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 299–314.

Solomon, I., Ariyo, A., & Tomassini, L. A. (1985). Contextual effects on the calibration of probabilistic judgments. *Journal of Applied Psychology*, *70*(3), 528–532.

Sonnemann, U., Camerer, C. F., Fox, C. R., & Langer, T. (2013). How psychological framing affects economic market prices in the lab and field. *Proceedings of the National Academy of Sciences*.

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G., & Burgman, M. (2010). Reducing overconfidence in the interval judgments of experts. *Risk Analysis*, *30*(3), 512–523.

Spellman, B. A., & Tenney, E. R. (2010). Credible testimony in and out of court. *Psychonomic Bulletin & Review*, *17*(2), 168–173.

Statman, M., Thorley, S., & Vorkink, K. (2006). Investor overconfidence and trading volume. *Review of Financial Studies*, *19*(4), 1531.

Stotz, O., & von Nitzsch, R. (2005). The Perception of Control and the Level of Overconfidence: Evidence from Analyst Earnings Estimates and Price Targets. *Journal of Behavioral Finance*, *6*(3), 121–128.

Subbotin, V. (1996). Outcome Feedback Effects on Under- and Overconfident Judgments (General Knowledge Tasks). *Organizational Behavior and Human Decision Processes*, *66*(3), 268–276.

Taylor, D., & McFatter, R. (2003). Cognitive performance after sleep deprivation: Does personality make a difference? *Personality and Individual Differences*, *34*(7), 1179–1193.

Teigen, K. H., & Jorgensen, M. (2005). When 90% confidence intervals are 50% certain: On the credibility of credible intervals. *Applied Cognitive Psychology*, *19*(4), 455–475.

Tenney, E. R. (2013). When Being Confident and Wrong Doesn't Matter For Credibility: People Evaluate Themselves and Others Differently. *Unpublished Manuscript*.

Tenney, E. R., Logg, J. M., & Moore, D. A. (2013). Optimistic About Optimism: The Belief That Optimism Improves Performance. *Unpublished Manuscript*.

Tenney, E. R., MacCoun, R. J., Spellman, B. A., & Hastie, R. (2007). Calibration trumps confidence as a basis for witness credibility. *Psychological Science*, *18*(1), 46–50.

Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: how young children and adults assess credibility. *Developmental Psychology*, *47*(4), 1065–77. doi:10.1037/a0023273

Tenney, E. R., & Spellman, B. A. (2011). Complex Social Consequences of Self-Knowledge. *Social Psychological and Personality Science* , *2* (4), 343–350. doi:10.1177/1948550610390965

Tenney, E. R., Spellman, B. A., & MacCoun, R. J. (2008). The benefits of knowing what you know (and what you don't): How calibration affects credibility. *Journal of Experimental Social Psychology*, *44*, 1368–1375.

Tetlock, P. E., & Mellers, B. A. (2011). Intelligent management of intelligence agencies: beyond accountability ping-pong. *The American Psychologist*, *66*(6), 542–54. doi:10.1037/a0024285

Thomas, J. P., & McFadyen, R. G. (1995). The confidence heuristic: A game-theoretic analysis. *Journal of Economic Psychology*, *16*(1), 97–113.

Thompson, L., & Loewenstein, G. (1992). Egocentric interpretations of fairness and interpersonal conflict. *Organizational Behavior and Human Decision Processes*, *51*(2), 176–197.

Tomassini, L. A., Solomon, I., Romney, M. B., & Krogstad, J. L. (1982). Calibration of auditors' probabilistic judgments: Some empirical evidence. *Organizational Behavior and Human Performance*, *30*(3), 391–406.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–31. doi:10.1126/science.185.4157.1124

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*(4), 547–567.

Tyszka, T., Zielonka, P., & Dubra, J. (2002). Expert Judgments: Financial Analysts versus Weather Forecasters Optimism and Overconfidence in Search. *Journal of Psychology and Financial Markets*, *3*(3), 152–160.

Vallone, R. P., Griffin, D. W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology*, *58*(4), 582–592.

Van Swol, L. M., & Sniezek, J. A. (2005). Factors affecting the acceptance of expert advice. *British Journal of Social Psychology*, *44*(3), 443–461.

Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, *93*(1), 1–13.

Yaniv, I., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General*, *124*(4), 424–32.

Yaniv, I., & Foster, D. P. (1997). Precision and accuracy of judgmental estimation. *Journal of Behavioral Decision Making*, *10*(1), 21–32.

Yates, J. F., Lee, J.-W., & Bush, J. G. (1997). General knowledge overconfidence: Cross-national variations, response style, and "reality ." *Organizational Behavior and Human Decision Processes*, *70*(2), 87–94.

Yates, J. F., Lee, J.-W., & Shinotsuka, H. (1996). Beliefs about overconfidence, including its cross-national variation. *Organizational Behavior & Human Decision Processes*, *65*(2), 138–147.