Organizational Behavior and Human Decision Processes 121 (2013) 246–255

Contents lists available at SciVerse ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

Cheap talk and credibility: The consequences of confidence and accuracy on advisor credibility and persuasiveness

Sunita Sah [a,b,*], Don A. Moore [c], Robert J. MacCoun [c]

[a] Georgetown University, McDonough School of Business, United States
[b] Harvard University, Edmond J. Safra Center for Ethics, United States
[c] University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 27 May 2011
Accepted 11 February 2013
Available online 26 March 2013
Accepted by Madan Pillutla

Keywords:
Judgment
Confidence
Advisors
Influence
Accuracy
Credibility
Calibration

ABSTRACT

Is it possible to increase one's influence simply by behaving more confidently? Prior research presents two competing hypotheses: (1) the confidence heuristic holds that more confidence increases credibility, and (2) the calibration hypothesis asserts that overconfidence will backfire when others find out. Study 1 reveals that, consistent with the calibration hypothesis, while accurate advisors benefit from displaying confidence, confident but inaccurate advisors receive low credibility ratings. However, Study 2 shows that when feedback on advisor accuracy is unavailable or costly, confident advisors hold sway regardless of accuracy. People also made less effort to determine the accuracy of confident advisors; interest in buying advisor performance data decreased as the advisor's confidence went up. These results add to our understanding of how advisor confidence, accuracy, and calibration influence others.

Introduction

We regularly rely on the advice of others. Businesses pay billions of dollars each year to receive advice from consultants. Patients rely on advice from their physicians. Individuals and businesses regularly seek financial advice. When making choices, even very personal ones, we take into account the opinions of others (Ajzen & Fishbein, 1980). In this paper, we ask how easy it is for advisors to manipulate their credibility or the persuasiveness of their advice by displaying more confidence than is justified.

When does confidence help and when does it hurt?

The research literature has offered two conflicting perspectives on the value of displaying confidence. The *confidence heuristic* maintains that people see confident advisors as more accurate, knowledgeable, and credible (Anderson, Brion, Moore, & Kennedy, 2012; Price & Stone, 2004). On the other hand, the *calibration hypothesis* asserts that advisors are more credible if they express confidence only when it is warranted, and that highly confident but inaccurate advisors lose credibility (Tenney, MacCoun, Spellman, & Hastie, 2007; Tenney, Spellman, & MacCoun, 2008).

Confident leaders can have more influence over others (Van Swol & Sniezek, 2005; Zarnoth & Sniezek, 1997), attain status more readily (Anderson et al., 2012), and are viewed as more competent (Anderson & Kilduff, 2009). Price and Stone (2004) argue that people use what they dubbed a 'confidence heuristic': People assume that more confident advice will be better, even when prior accuracy information suggests it wasn't always so. These results raise the question of whether expressing confidence always benefits a leader. Can advisors use strategically expressed confidence as a means of influence (Yates, Price, Lee, & Ramirez, 1996)? What happens when confident people are wrong? Can confidence backfire?

Tenney and her colleagues (Tenney et al., 2007, 2008) demonstrate that people attend to more than simple confidence. They also attend to *calibration*. In other words, advisors are perceived as credible if they express confidence only when it is warranted. Tenney et al. (2007, 2008) gave their participants hypothetical examples of eyewitness testimony. Their participants reported that errors did a great deal of damage to confident witnesses' credibility. By contrast, the credibility of the less confident witness is not so severely undermined when he is found to be incorrect. This suggests that calibration is more important than confidence.

* Corresponding author. Address: Georgetown University, McDonough School of Business, Office 575, 37th and O Street, Rafik B. Hariri Building, Washington, DC 20057, United States.
E-mail address: ss3250@georgetown.edu (S. Sah).

0749-5978/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.obhdp.2013.02.001

Confidence, accuracy, and calibration

Confidence and accuracy contribute to an advisor's credibility or influence and prior studies differ in terms of which of these factors matter and whether they are additive or interact (Berman & Cutler, 1996; Brewer & Burke, 2002; Tenney et al., 2007, 2008). Earlier studies examining naturally-occurring confidence found that confidence increased persuasiveness (Phillips, 1999; Sniezek & Buckley, 1995; Sniezek & Van Swol, 2001; Van Swol & Sniezek, 2005; Zarnoth & Sniezek, 1997). Confidence was often correlated with accuracy in these studies; therefore, advisees may have been rational to take confidence as a cue to expertise. In other words, confidence can be an important source of information (Bonaccio & Dalal, 2006; Yaniv, 1997). Confidence and accuracy often co-vary (Sniezek, 1992) but the relationship can be weak (Deffenbacher, 1980; Kassin, 1985; Klayman, Soll, González-Vallejo, & Barlas, 1999; Shaw & McClure, 1996; Wells, Lindsay, & Ferguson, 1979; Wells, 1993) and sometimes they are uncorrelated (Brewer & Wells, 2006; Ekman & O'Sullivan, 1991).

Like Price and Stone (2004) and Tenney et al. (2007, 2008), we manipulate confidence independent of accuracy to avoid the potential confound inherent in research designs that examine naturally occurring expressions of confidence. Then we seek to resolve the discrepancy between the findings from Price and Stone and Tenney et al. using a novel experimental paradigm that includes real incentives for accuracy. We suspect that the ease of calibration played a role in these differing results with participants finding it easier and less effortful to calibrate in Tenney et al.'s (2007, 2008) studies than in Price and Stone (2004). We seek to clarify the role of clear feedback on an advisor's performance in determining when people strive to calibrate and when they use the confidence heuristic.

Tenney et al. (2007) describe good calibration as "being confident when right and unconfident when wrong" (2007, p. 47), and claim that "when people get evidence about an informant's calibration (i.e., her confidence–accuracy relationship) they override reliance on confidence or accuracy alone" (Tenney et al., 2008, p. 1368; see also, Chaiken & Eagly, 1976; Tenney & Spellman, 2011). In order for calibration to occur, accuracy information must be readily available. This is the calibration hypothesis:

Hypothesis 1. In the presence of clear accuracy feedback, calibration will be more important than confidence.

The moderating role of feedback

In previous tests of the calibration hypothesis (Tenney et al., 2007, 2008), relevant calibration evidence was readily available. Study 1 explores the impact of advisor confidence, accuracy, and calibration on perceived credibility and actual influence when advisor accuracy is freely available and clear. But in everyday settings, performance data is often either unavailable or costly to obtain and interpret. Study 2 further explores boundaries of the calibration hypothesis when information on accuracy is not available or is costly to obtain.

According to Tenney et al.'s (2008, p. 1369) "presumption of calibration" hypothesis, "people initially presume, in the absence of relevant evidence, that informants are well calibrated," giving an advantage to more confident actors. But "people will override that initial presumption when evidence that enables the assessment of the informant's calibration becomes available," at which point it is good calibration (confidence that matches accuracy) rather than high confidence that will make the source credible. Note that the "presumption of calibration" hypothesis suggests a two-stage process: First, equating high confidence with high accuracy, perhaps relying on normative rules of reasoning, i.e., a deliberative process (Evans, 2003), and second, once accuracy information is available, calibrating accordingly. In contrast, the confidence heuristic implies an effortless intuitive process (Kahneman, 2003; Masicampo & Baumeister, 2008; Tversky & Kahneman, 1974) in which confidence is used as a peripheral cue. Accuracy information may, or may not, be present but confidence is used as a short-cut for advisor credibility and influence. It's possible that the first stage in the presumption of calibration hypothesis is also intuitive and utilizes the effortless confidence heuristic; however, the key difference is that once accuracy information is available, according to the calibration hypothesis, people will use it to calibrate their advisors. In the absence of accuracy information, either of these two accounts, the presumption of calibration or the confidence heuristic, will lead to the same outcome: Greater confidence results in greater credibility and influence.

Hypothesis 2. When feedback on accuracy is unavailable, people will assign greater credibility to, and be more persuaded by, confident advisors than their low confidence counterpart.

Study 2 also investigates the situation in which feedback is potentially available but costly. We predict, as in so many other domains (Payne, Bettman, & Johnson, 1990; Smith, Mitchell, & Beach, 1982), that many people will revert to a simple heuristic—in this case, the confidence heuristic—as the cost of deliberation increases. Although we expect some people to seek calibration information, we also expect that many advisees will simply rely on the confidence heuristic.

Hypothesis 3. When accuracy feedback is costly, many advisees will revert to the confidence heuristic and find high confidence credible and persuasive.

Our test of Hypothesis 3 also provides an opportunity to examine whether advisor confidence influences advisees' decisions about purchasing or seeking feedback on advisor performance.

Our experimental paradigm

Previous researchers have investigated the influence of confidence when feedback was unavailable (for example, Sniezek & Van Swol, 2001), however, those studies employed naturally-occurring confidence in which confidence and accuracy happened to be positively correlated, thus complicating causal inference. Advisors in the Sniezek and Van Swol experiments were also rewarded when advisees were correct. Therefore advisors wanted to calibrate their confidence to their accuracy and both parties knew it. However, many professional advisors are rewarded when people buy or follow their advice, regardless of whether it is in the recipients' best interests and regardless of whether it is accurate (Sah & Loewenstein, 2012; Van Swol, 2009). Our paradigm systematically manipulates confidence and can therefore investigate what happens when confidence is not a valid cue to accuracy.

We also introduce two other methodological contributions. First, our stimulus materials employ a continuous outcome metric. By contrast, Price and Stone's (2004) advisors predicted whether a stock would go up or down in value, and witnesses in Tenney et al.'s stimulus cases testified that a suspect was present or absent at the scene of a crime. The benefit of using a continuous outcome metric is threefold:

(1) It allows us to disentangle proximity to the correct answer from confidence in one's answer. This distinction is helpful, for instance, in testing Hypothesis 1 (the calibration hypothesis), regarding advisors who are low in confidence but whose answers are nevertheless close to the correct answer.

This is not possible with dichotomous outcomes, where being close to the actual outcome necessarily requires an extreme estimate of probability. For example, say the question is, "What is the probability that the person in the picture weighs more than 150 lb?" If, in fact, the person weighs 152 lb, then the judge who responds with 100% confidence is more correct than the one who is only 51% sure. But for continuous judgments, I can be close to the right answer, guessing that the person weighs 151 lb, but still lack confidence, claiming to be only 30% sure that my guess is within 5 lb of the truth. Instances in which an advisor is high in accuracy but low in confidence are essential for independently manipulating those key variables, as we do.

(2) It avoids some of the confusion inherent in half-scale measures of confidence produced by dichotomous outcomes (Hoffrage, 2004), where a 50% rating implies complete lack of knowledge rather than moderate confidence (Fischhoff & Bruine De Bruin, 1999). It can feel wrong to claim 50% confidence when you are completely clueless. But to express 0% confidence about a dichotomous judgment implies one is 100% sure that the other option is correct.

(3) As Keren and Teigen (2001) point out, preference for extreme confidence may be driven by the fact that it makes for sharper contrasts and therefore clearer discriminations, especially with dichotomous outcomes. This is less of an issue with continuous scale judgments, where people can be extremely confident regarding the accuracy of estimates of middling size. If the question is, "Will next year be warmer than this year?" then advisors who wish to distinguish themselves must make extreme confidence judgments. A continuous judgment would allow an advisor to express certainty that next year will be no warmer or colder than this year, forecasting zero change with high confidence.

Second, we include a behavioral measure of the extent of the advisor's actual influence to supplement ratings of perceived credibility. Although one would expect perceived credibility and actual influence to be positively correlated, the actual empirical relationship between the two measures of effectiveness has been little studied and is likely to be far from 1.0 (Dillard, Weber, & Vail, 2007). Perhaps this is because people are sometimes not consciously aware of how they are influenced by a stimulus (Nisbett & Wilson, 1977). One benefit of utilizing behavioral measures is that they can often reveal response patterns of which respondents themselves are unaware of. It is possible, for instance, that people report explicitly disliking the confident but potentially inaccurate advisor, even while they are more persuaded by the confident advice. Our perceived credibility ratings are explicit cognitions, but because actual influence can reflect both implicit and explicit influences, these two measures may give different results (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Implicit processing may be more likely to rely on simple heuristics that people are often unaware of. While actual influence is more consequential than perceived influence, the latter is also of interest, because perceived influence affects the advisor's reputation and prestige – perhaps even more than actual influence.

In sum, Study 1 investigates the effects of advisor confidence, accuracy and calibration in the presence of full costless feedback. We conduct a conceptual replication of Tenney et al. (2007, 2008), exogenously manipulating confidence and accuracy in a new continuous judgment paradigm (with 5 confidence–accuracy pairings), and using two different dependent variables (perceived advisor credibility and actual behavioral persuasiveness). We test whether the calibration or confidence heuristic hypothesis better accounts for the data. Study 2 examines the role of feedback as a key moderator and looks for evidence of reversion to heuristics when feedback is unavailable or costly—a potential boundary effect to the calibration hypothesis.

Study 1

We selected our experimental task—estimating how much people weigh by examining their photographs—for three reasons: (1) It is easily understood by participants; (2) it is relatively engaging; and (3) it comes from a domain where there are likely to be individual differences in accuracy. Our studies employ a simple form of advice: a numerical estimate from an advisor. As Yaniv (2004) explains, "Simple as it is, numerical advice has an important function in individual as well as organizational decisions. Physicians, weather forecasters, genetic consultants, and lawyers, just to name a few, are all in the business of communicating their forecasts and uncertain estimates to others facing decisions" (p. 2).

Method

Participants. One hundred eighty-four individuals (over 90% students, 49% male, 92% between 18 and 35 years of age), recruited using an advertisement on the website of an eastern United States research university, completed the study. Participants were told that their chances of winning one of several $100 prizes increased with their performance on the task, and that there would be one prize awarded for every 20 participants.

Experimental task. The experimental procedure was conducted on computers in the university lab. Participants were asked to estimate the weights of 10 different people from photographs. They were told (truthfully) that the people in the photographs varied in weight between 90 and 230 lb. Participants assigned probabilities to 14 ten-pound intervals from 90 to 230 lb corresponding to their estimate of the person's weight. The photographs they saw only included the target person's head and shoulders as shown in Fig. 1.

For each photo, participants first made an initial estimate. They could enter their estimate (probability that the weight fell into that 10-pound interval) in one interval (if they were 100% confident that the weight of the person in the photograph fell into that interval) or they could distribute their confidence across intervals if they were less sure. They then received advice which consisted of one (the highest) confidence rating from the advisor for a single 10-pound weight band (e.g., "*Your advisor is 65% confident that the person's weight is between 90–99 lbs*"). Participants were told that their advisors had seen a full-body picture of each photographed person and thus had better information than they did. After receiving the advice, participants had the opportunity to revise their initial estimates, therefore participants provided pre- and post-advice weight estimates for each of 10 photographs. After submitting their revised estimates, participants learned the photographed person's actual weight in the accuracy phase (more details on the phases below). Participants were rewarded for their accuracy using a version of the quadratic scoring rule (see Moore & Healy, 2008).[1] They could earn up to a maximum of two lottery tickets per photograph.

Advisors, design and procedure. Each participant was randomly assigned to receive advice from one of four fictional advisors, representing a 2 (confidence: high vs. low) × 2 (accuracy: high vs. low) between-subject factorial design. This gives us two well-calibrated advisors (high accuracy/high confidence and low accuracy/low confidence) and two poorly-calibrated advisors (low

[1] Participants earned $2p$ lottery tickets for the correct interval (where p equals the probability assigned to the interval), minus the summed total of p^2 lottery tickets for each of the other (incorrect) intervals.

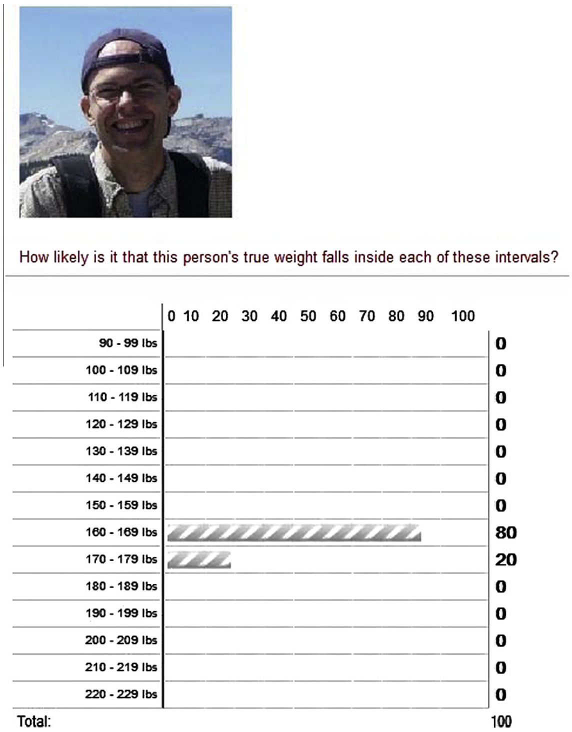


Fig. 1. Example weight estimate screen. The participant has indicated an 80% probability that the weight of the person in the photograph is between 160–169 lbs and a 20% probability that the weight is between 170 and 179 lbs.

accuracy/high confidence and high accuracy/low confidence).[2] High confidence advisors gave confidence ratings between 85% and 100% with a mean of 95%, and low confidence advisors varied between 25% and 35% confidence with a mean of 30%. We selected these confidence levels based in part on pre-test results showing that this is a plausible range for the level of confidence people express on this task.

High accuracy advisors gave correct advice and low accuracy advisors gave incorrect advice for the first five consecutive photographs. Since we use a continuous rather than a dichotomous judgment, it is realistic that an advisor (even a highly confident one) could be wrong in five successive trials. Our levels of inaccuracy vary within the five inaccurate trials, ranging from two to four ten-pound intervals above or below the correct answer (and thus were not biased in any one direction). Prior research shows that people routinely err this frequently at this task, even when they are motivated to do well (Gino & Moore, 2007; Moore & Klein, 2008). After this first "accuracy phase," participants saw an additional five photographs in which the advice (which was always correct) indicated only a weight interval but no confidence rating, and participants got no feedback regarding the correct weight. We designed this "test phase" to examine the degree to which an advisor holds sway, allowing us to test the durability of effects such as advisor confidence and accuracy on persuasiveness. At the end of the 10 photographs, participants rated how much they agreed or disagreed with the following statements used to assess advisor credibility: "My advisor was competent," "I trusted my advisor," "I liked my advisor," "I took my advisor's advice," "My advisor was consistent," and "My advisor was reliable."

Dependent Measures. We employed two different dependent measures:

[2] Like Tenney et al. (2007, 2008), we use "calibration" as the general term to refer to the confidence–accuracy relation: Well calibrated advisors have a gamma correlation of 1 and poor-calibrated advisors have a gamma correlation of -1.

Perceived advisor credibility was determined by averaging the six advisor ratings.

Advisor persuasiveness was measured with the 'Weight of Advice' measure used by Yaniv (2004), which measures the amount that participants change their estimates after getting advice. It provides a behavioral measure of persuasiveness. To get a "best guess" point estimate from each subject's reported belief distribution, we multiplied the percentage confidence attached to each ten-pound interval with its relevant weight in pounds (e.g., the mid-point of the interval range), and summed across all 14 intervals. Then we calculated the absolute difference from participants' pre-advice estimate to their post-advice estimate. We divided this by the absolute difference of participants' pre-advice estimate from the advisor recommendation. The formula is:

$$\text{Persuasiveness} = \frac{|\text{Advisee final estimate} - \text{Advisee initial estimate}|}{|\text{Advisor recommendation} - \text{Advisee initial estimate}|}$$

Results

Manipulation checks. The manipulations worked as intended. Participants with high confidence advisors rated their advisors as more confident ($M = 4.0$, on a scale of 1–5 where 1 is low and 5 is high) than participants with low confidence advisors ($M = 2.9$), $F_{(1,180)} = 51.63$, $p < .001$. High or low advisor accuracy was also correctly identified by participants; participants with high accuracy advisors rated their advisors as more accurate ($M = 4.1$) than those with low accuracy advisors ($M = 2.3$), $F_{(1,180)} = 183.48$, $p < .001$.

The correlation between perceived advisor credibility and persuasiveness was positive and significant, $r = .44$, $p < .01$.

Perceived advisor credibility. We asked participants to rate their advisors on six dimensions. We factor-analyzed the answers to these questions, and a single factor explained approximately 74% of the variance. Therefore, we averaged them to compute an average advisor rating (Cronbach's alpha = .93) of perceived advisor credibility.

A 2 (advisor confidence: high vs. low) $\times$ 2 (advisor accuracy: high vs. low) between-subjects ANOVA for advisor credibility revealed that high accuracy advisors were rated significantly higher ($M = 3.84$) than low accuracy advisors ($M = 2.53$), $F_{(1,180)} = 142.03$, $p < .001$. There was no significant effect for confidence, (M(high confidence) = 3.11 versus M(low confidence) = 3.26), $F_{(1,180)} = 1.70$, $p = .19$, however, there was a significant interaction between advisor confidence and accuracy, $F_{(1,180)} = 18.54$, $p < .001$. The simple effect of confidence was significant for high accuracy and low accuracy advisors, although in opposite directions: For high accuracy advisors, high confidence was rated higher ($M = 4.00$) than low confidence ($M = 3.67$), $F_{(1,180)} = 5.09$, $p = .026$. However, for low accuracy advisors, high confidence was rated significantly lower ($M = 2.23$) than low confidence ($M = 2.84$), $F_{(1,180)} = 14.29$, $p < .001$ as illustrated in Fig. 2A. This result is important, as it supports the calibration hypothesis over the confidence heuristic (support for *Hypothesis 1*).

Advisor persuasiveness. Persuasiveness had a mean of .71. The higher the persuasiveness score, the more a participant moved his or her estimate from pre- to post-advice. Participants with a persuasiveness score of 1 moved their estimates to match their advisors' recommendations. A 2 (advisor confidence: high vs. low) $\times$ 2 (advisor accuracy: high vs. low) ANOVA revealed that high accuracy advisors were more persuasive ($M = .84$) than low accuracy advisors ($M = .54$), $F_{(1,180)} = 32.53$, $p < .001$. The main effect for confidence was not significant (M(high confidence) = .67 versus M(low confidence) = .71), $F_{(1,\ 180)} = .71$, $p = .40$. Unlike the advisor credibility rating, there was no significant interaction for our behavioral measure of persuasiveness, $F_{(1,180)} = 1.60$, $p = .21$, see Fig. 2B.[3]

Discussion

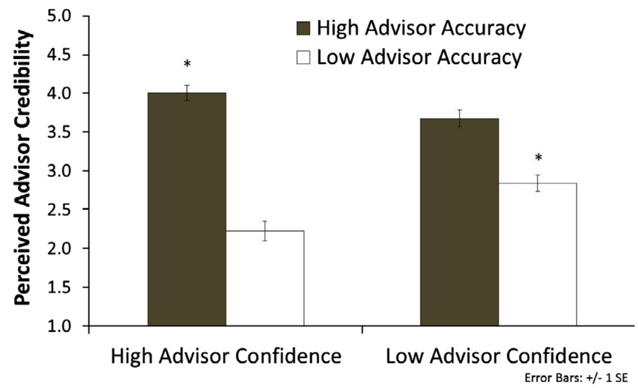
Perceived advisor credibility. Advisor ratings reveal that accuracy moderates the effect of confidence. Highly accurate advisors benefit from confidence, whereas inaccurate advisors prove more credible when they express less confidence. Highly confident but inaccurate advisors received the lowest ratings of credibility. This supports Tenney et al.'s (2008) calibration hypothesis (and our *Hypothesis 1*) in a new research paradigm.

The modest (low confidence but accurate) advisors also received relatively high credibility ratings but still significantly lower than the highly confident accurate advisors; therefore, although accuracy was clearly an important component in determining credibility (note the large *F*-statistic), it appears that low confidence accurate advisors were punished somewhat for low calibration. It is possible that advisees could have attributed some of their advisor's accuracy to luck (rather than modesty, for example) since low confidence indicates advisors were not sure of the answer and may have just guessed correctly.

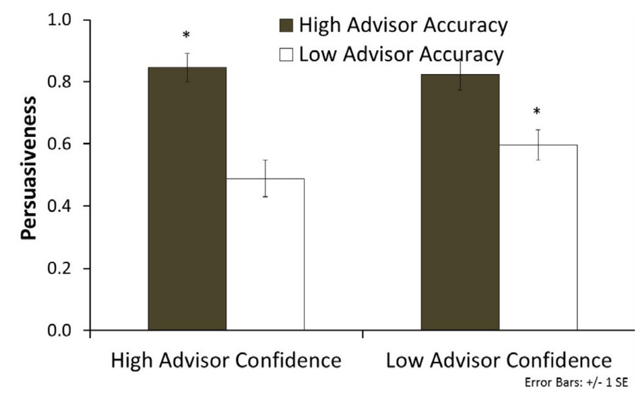
Advisor persuasiveness. Our behavioral measure of persuasiveness did not incorporate advisor confidence. Whereas perceived advisor credibility was influenced by both confidence and accuracy from advisors, actual influence (persuasiveness) was largely driven by accuracy.

We did not expect perceived credibility and persuasion to be perfectly correlated (Dillard et al., 2007) and we see three potential reasons why our results differ for the two dependent measures. First, it is possible that our persuasiveness measure is simply lower in reliability than the credibility composite. Second, while our credibility ratings tap explicit cognition, the actual persuasiveness index may also tap implicit processes less amenable to self-report. And third, credibility and persuasiveness are related, but are conceptually and empirically distinct constructs that may well differ in their determinants (Ajzen & Fishbein, 1977; Dillard et al., 2007). Both of our measures found advisor accuracy to be a highly important component–understandably, you want to follow your advisor when she is correct but perceptions of credibility may incorporate advisor confidence.

Overall, our results from Study 1 provide little support for the confidence heuristic in a situation where accuracy information is clear and available. Advisees used accuracy information wisely in updating their estimates. Advisees also took confidence into account when rating their advisor's credibility, consistent with the calibration hypothesis. In Study 2, we test for further boundary conditions of the calibration hypothesis by examining the relative importance of accuracy, confidence, and calibration when the availability and cost of accuracy information vary.



Panel A: Perceived Advisor Credibility



Panel B: Persuasiveness

* Well-calibrated advisors (high confidence when accurate and low confidence when inaccurate)

Fig. 2. Study 1.

Study 2

According to the logic of the calibration account, advisor confidence should backfire when the respondent actually has an opportunity to assess the advisor's calibration. This was relatively easy for them to do in experiments by Tenney et al. (2007, 2008) and in our Study 1. Often, however, assessing accuracy is more difficult than that. Therefore, Study 2 examines whether people revert to the simple confidence heuristic when accuracy is more difficult to assess.

Study 2's paradigm was similar to that of Study 1, except for its three-level manipulation of accuracy feedback about the advisor: free feedback vs. no feedback vs. costly feedback. The free feedback condition was similar to Study 1, and we expected to replicate Study 1's findings. In the no-feedback condition, advisees got no information about advisor accuracy. Hypothesis 2 predicts that people will assign greater credibility to, and be more persuaded by, the more confident advisor, which is consistent with both the confidence heuristic and the presumption of calibration hypothesis.

In the costly feedback condition, feedback was available, but only for advisees sufficiently motivated to pay for it. We predicted that many people would revert to the confidence heuristic as the cost of deliberation increased and, therefore, would be more persuaded by high confidence advisors than low confidence advisors (*Hypothesis 3*). We also investigated when people choose to buy feedback.

Method

Participants. Three hundred and seventy-seven individuals (over 90% students, 42% males, 92% between 18–35 years of age) from the online university subject pool of a large research university in the western U.S. participated for a chance at winning a $100 prize.

Experimental task and procedure. As in Study 1, the participants' task was to estimate the weight of different people from photographs. This time there were only five photographs (we eliminated the test phase, as described below) and participants were again rewarded for accuracy in the same way, earning up to two lottery tickets per photograph.

[3] We repeated this analysis using data from the first 5 photographs-the accuracy phase (in which participants are given a confidence rating and feedback on the advisor's accuracy immediately after entering their final estimate). We find the same results—a significant main effect for advisor accuracy, $M(\text{high accuracy}) = .85$ vs. $M(\text{low accuracy}) = .50$, $F_{(1,180)} = 13.40$, $p < .001$, and no significant interaction or main effect for confidence.

As in Study 1, the experimental design varied both advisor confidence and accuracy. This study adds the feedback manipulation, producing a 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high vs. low) × 3 (feedback: free vs. none vs. costly) between-subject design. As before, participants with high accuracy advisors received correct advice for five consecutive photographs, and low accuracy advisors gave incorrect advice for five consecutive photos. Advisor ratings were completed after these five photographs.

In the free feedback condition ($n = 120$), participants received feedback on the advisor's accuracy (after the participant entered their final estimate on that photograph), such as, "Your advisor was correct," if the advisor gave an estimate in the correct weight interval, or, "Your advisor was 3 intervals away from the correct answer," if the advisor gave an estimate three intervals above or below the correct interval. In the no-feedback condition ($n = 131$), participants did not receive any information on the accuracy of their advisor. In the costly feedback condition ($n = 126$), participants could opt to receive feedback. Before each photograph, participants were asked, "*Would you like to buy feedback on your advisor's accuracy for the next photograph? This will cost you 1 lottery ticket*." The feedback was given in the same format as in the free feedback condition, after the participant gave their final decision. Allowing participants to buy feedback on each photo allows participants to experience advisor confidence (and possibly advisor accuracy) before deciding whether to buy or not on subsequent rounds.

We eliminated the test phase in this study, and examined responses within the accuracy phase. The test phase in Study 1 had advice with no confidence ratings or accuracy information. Due to Study 2's manipulation of the availability of accuracy information, including the test phase would have been problematic.[4] We examined the same two dependent measures as in Study 1: perceived advisor credibility (Cronbach's alpha = .91) and advisor persuasiveness.

Results

Manipulation checks. The manipulations worked as intended. High confidence advisors were rated as more confident ($M = 4.1$) than low confidence advisors ($M = 3.3$), $F_{(1,371)} = 53.03$, $p < .001$. The feedback manipulation moderated participants' sense of their advisor's confidence, resulting in a significant interaction, $F_{(2,371)} = 6.67$, $p = .001$, with participants rating high confidence advisors similarly regardless of feedback condition, M(free feedback) = 4.1, M(no feedback) = 4.2, M(costly feedback) = 3.9, $F_{(1,371)} = 1.57$, $p = .21$, but rating low confidence advisors in the no-feedback condition as the least confident, M(no feedback) = 3.0, compared to low confidence advisors in the free feedback ($M = 3.4$) and costly feedback ($M = 3.6$) conditions, $F_{(1,371)} = 13.35$, $p < .001$.

Participants with high accuracy advisors rated their advisors as more accurate ($M = 3.8$) than those with low accuracy advisors ($M = 2.6$), $F_{(1,371)} = 172.96$, $p < .001$. As expected, the more difficult it was to obtain accuracy feedback, the less able participants were to distinguish between accurate and inaccurate advice, as revealed by a significant accuracy × feedback interaction, $F_{(2,371)} = 95.50$, $p < .001$; participants in the free feedback condition were most able to distinguish between high ($M = 4.8$) and low accuracy ($M = 2.0$) advisors, $F_{(1,371)} = 336.84$, $p < .001$; those in the costly feedback condition could also distinguish between high ($M = 3.4$) and low accuracy ($M = 2.8$) advisors, $F_{(1,371)} = 16.28$, $p < .001$; however, those in the no-feedback condition were unable to distinguish between high ($M = 3.2$) and low accuracy ($M = 3.2$) advisors, $F_{(1,371)} = .001$, $p = .98$.

The correlation between perceived advisor credibility and persuasiveness was positive and significant, $r = .20$, $p < .01$. We have separated the results in the following sections to highlight the effect of the feedback manipulation.

Free feedback condition

Perceived advisor credibility. Similar to Study 1, a 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high vs. low) ANOVA revealed a highly significant main effect for accuracy, M(high accuracy) = 4.34 versus M(low accuracy) = 2.69, $F_{(1,116)} = 152.11$, $p < .001$. The main effect for confidence was not significant, M(high confidence) = 3.40 versus M(low confidence) = 3.62, $F_{(1,\ 116)} = 2.76$, $p = .10$. Again, there was a significant interaction, $F_{(1,116)} = 7.66$, $p = .007$, as shown in Fig. 3A. High confidence low-accuracy advisors were rated the lowest ($M = 2.39$), significantly below low confidence low-accuracy advisors ($M = 2.98$), $F_{(1,116)} = 10.15$, $p = .002$, again demonstrating support for the calibration hypothesis (*Hypothesis 1*).

This time, we note that underconfidence was tolerated in accurate advisors: high confidence high-accuracy advisors were not rated significantly higher ($M = 4.41$) than low confidence high-accuracy advisors ($M = 4.27$), $F_{(1,116)} = .59$, $p = .44$.

Advisor persuasiveness. A 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high vs. low) ANOVA revealed, as in Study 1, that high accuracy was more persuasive ($M = .82$) than low accuracy ($M = .53$), $F_{(1,116)} = 14.21$, $p < .001$ (Fig. 3B).[5] The effect of confidence did not attain significance (M(high confidence) = .72 versus M(low confidence) = .63), $F_{(1,116)} = 1.64$, $p = .20$, nor did it interact with accuracy, $F_{(1,116)} = .28$, $p = .60$.

No feedback condition

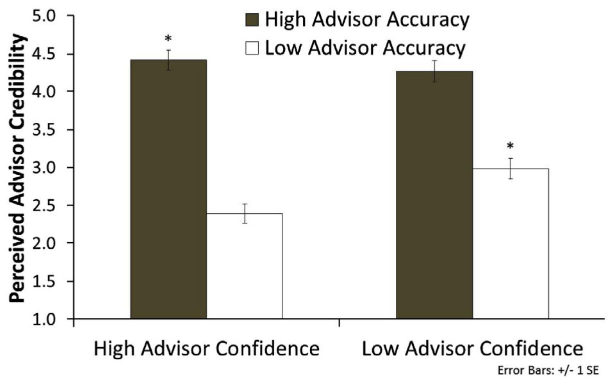
Perceived advisor credibility. The 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high vs. low) ANOVA revealed only a significant main effect for confidence such that high confidence advisors were rated higher ($M = 3.54$) than low confidence advisors ($M = 3.20$), $F_{(1,127)} = 7.82$, $p = .006$, giving support to *Hypothesis 2*. The main effect for accuracy was not significant, M(high accuracy) = 3.34 versus M(low accuracy) = 3.40, $F_{(1,127)} = .18$, $p = .67$, nor was there an interaction, $F_{(1,127)} = .11$, $p = .74$.

The presumption of calibration hypothesis suggests that people will assume, in the absence of other information, that advisors are well-calibrated so that high confidence advisors will be deemed as accurate, and thus credible. However, when comparing the high confidence accurate advisor in the no-feedback condition with the same advisor (high confidence and accurate) in the free feedback condition, it is interesting to note that advisor credibility is significantly lower in the no-feedback condition ($M = 3.54$) compared to the feedback condition ($M = 4.41$), $F_{(1,62)} = 37.36$, $p < .001$. Although high confidence advisors are deemed more credible than low confidence advisors when there is no performance data, these advisors do not get the full boost of credibility that they should from the presumption of high accuracy (calibration hypothesis), since high confidence advisors who are shown to be accurate have significantly higher credibility ratings. This suggests that any presumption of calibration is partial or tentative. Since this measure captures explicit cognitions, the reluctance to assume full calibration may be deliberate.

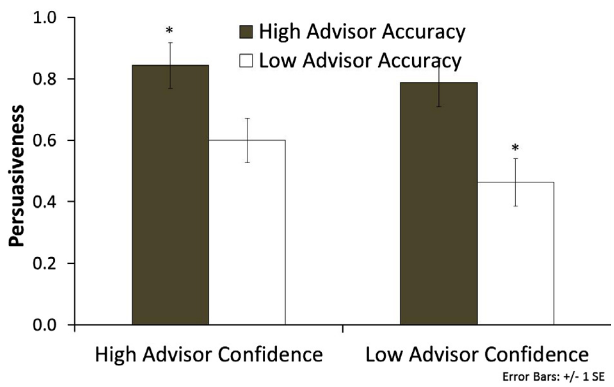
Advisor persuasiveness. The results here somewhat parallel the results for perceived advisor credibility and support *Hypothesis 2*: participants with high confidence advisors were more persuaded ($M = .93$) than those with low confidence advisors ($M = .60$), $F_{(1,127)} = 9.05$, $p = .003$. The main effect for accuracy was not

[4] Since, the persuasiveness results from Study 1 were similar for the "accuracy" and "test" phases, we hoped this would not alter our results substantially.

[5] We replicated the persuasiveness results from Study 1, this time without the test phase.



Panel A: Perceived Advisor Credibility



Panel B: Persuasiveness

*Well-calibrated advisors (high confidence when accurate and low confidence when inaccurate)

Fig. 3. Study 2. Free feedback condition.

significant, M(high accuracy) = .77 versus M(low accuracy) = .79, $F_{(1,\ 127)} < .01$, $p = .99$, nor was there an interaction, $F_{(1,127)} = .45$, $p = .51$.

This time with our behavioral measure, there was no significant difference between high confidence accurate advisors in the no-feedback condition ($M = .89$) compared to the free feedback condition ($M = .84$), $F_{(1,62)} = .12$, $p = .73$, and thus advisees appear to be using confidence as a full proxy for accuracy. Since this measure captures implicit cognitions, this may suggest a greater reliance on the confidence heuristic than the explicit credibility ratings demonstrated.

Costly feedback condition

Feedback. Approximately 44% of the participants bought feedback at least once in this condition. Overall, participants bought feedback approximately 22% of the time it was offered. Measuring the number of times feedback was bought (from 0 to 5), we find that participants bought more feedback when the advisor displayed low confidence ($M = 1.58$) compared to high confidence ($M = .63$), $F_{(1,122)} = 11.47$, $p = .001$. There was no effect of advisor accuracy on the number of times feedback was bought, $F_{(1,122)} = .001$, $p = .97$, nor an interaction of advisor accuracy and confidence, $F_{(1,122)} = .10$, $p = .75$. When advisors exhibited low confidence, 53% of participants bought feedback at least once, whereas when advisors had high confidence, only 35% of participants bought feedback, $\chi^2(1, N = 126) = 4.37$, $p = .037$.

Perceived advisor credibility. A 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high vs. low) ANOVA revealed that high accuracy advisors were rated higher ($M = 3.35$) than low accuracy advisors ($M = 3.06$), $F_{(1,122)} = 4.50$, $p = .036$. The main effect for confidence was not significant, $F_{(1,122)} = 1.40$, $p = .24$ nor was there an interaction, $F_{(1,122)} = .48$, $p = .49$.

We also conducted a 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high vs. low) × 2 (bought feedback: yes vs. no)[6] between-subjects ANOVA for advisor credibility ratings. Again, accurate advisors were perceived as more credible than inaccurate advisors, $F_{(1,118)} = 10.91$, $p = .001$. Participants who bought feedback rated their advisors as more credible ($M = 3.46$) than those participants who did not buy feedback ($M = 3.07$), $F_{(1,118)} = 8.98$, $p = .003$. There was also a significant interaction between accuracy and buying feedback, $F_{(1,118)} = 21.72$, $p < .001$, in that those who bought feedback and had accurate advisors rated their advisors as more credible ($M = 3.96$) than those who bought feedback and had inaccurate advisors ($M = 2.94$), $F_{(1,122)} = 31.03$, $p < .001$. Those who did not buy feedback rated their advisors as less credible regardless of accuracy, M(high accuracy) = 2.98 versus M(low accuracy) = 2.94, $F_{(1,122)} = 1.33$, $p = .25$, as shown in Fig. 4A.

Advisor persuasiveness. A 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high versus low) ANOVA revealed that high confidence advisors were more persuasive ($M = .77$) than low confidence advisors ($M = .54$), $F_{(1,122)} = 6.59$, $p = .011$ (supporting *Hypothesis 3*). The main effect for accuracy was not significant, $F_{(1,122)} = 1.88$, $p = .17$, nor was the interaction, $F_{(1,122)} = 3.46$, $p = .065$.

Again, we conducted a 2 (advisor confidence: high vs. low) × 2 (advisor accuracy: high vs. low) × 2 (feedback bought: yes vs. no) ANOVA for advisor persuasiveness. Participants with high confidence advisors were more persuaded than those with low confidence advisors, $F_{(1,\ 118)} = 6.61$, $p = .01$. There were also two significant interactions: one between accuracy and buying feedback, $F_{(1,118)} = 7.29$, $p = .008$, and the other between confidence and buying feedback, $F_{(1,118)} = 3.82$, $p = .05$. As with our perceived advisor credibility ratings, the accuracy and buying feedback interaction revealed that participants who bought feedback and had accurate advisors were more persuaded ($M = .91$) than those who bought feedback and had inaccurate advisors ($M = .50$), $F_{(1,122)} = 9.49$, $p = .003$. Those who did not buy feedback were also less persuaded regardless of accuracy, M(high accuracy) = .58 versus. M(low accuracy) = .75, $F_{(1,122)} = 2.31$, $p = .13$, see Fig. 4B. The confidence and buying feedback interaction revealed that participants who did not buy feedback were more likely to be persuaded by high confidence ($M = .83$) compared to low confidence ($M = .42$), $F_{(1,122)} = 11.98$, $p = .001$, but there was no effect of confidence among those who bought feedback, M(high confidence) = .67 versus M(low confidence) = .68, $F = .01$, $p = .92$, see Fig. 4B.

The perceived advisor credibility ratings were significantly lower among those advisees who did not buy feedback. Due to the low credibility they assign these advisors, plausible interpretations are that advisees did not bother buying feedback to check accuracy because they were not intending to use the advisor's advice, or perhaps advisees were being even more cautious regarding the presumption of calibration in advisors with unknown accuracy. The persuasiveness measure, however, indicates that those who do not buy feedback were influenced by confidence (*see* Fig. 4, A and B). This result can be interpreted as conditional support for the use of a confidence heuristic (captured by our implicit measure).

Discussion

[6] For some of our analyses in the costly feedback condition, we include whether feedback was purchased as an independent variable, but we caution that because participants self-select their level of this variable, "effects" that include it are more causally ambiguous than effects that only involve the randomly assigned confidence and accuracy variables.

When feedback was free and easily accessible, we mainly replicated the findings of our first study: overconfident advisors received low credibility ratings (support for calibration and *Hypothesis 1*) and advisor accuracy determined persuasiveness. This time, perceived credibility ratings (advisors were rated straight after the five confidence–accuracy pairings) revealed that underconfidence or modesty was tolerated in accurate advisors, i.e., they (unlike the inaccurate but confident advisors) were not punished for low calibration. This suggests that, although calibration was a factor in determining credibility, accuracy appears to be the most important component in determining both credibility and actual influence.

In the absence of feedback, participants relied on advisor confidence—rating high confidence advisors as significantly more credible and being more persuaded by them. Although our credibility ratings showed a significant effect for high confidence, high confidence advisors who were known to be accurate (seen in the free feedback condition) received significantly higher credibility ratings than high confidence advisors with unknown accuracy. This suggests that the presumption of calibration was tentative and confident advisors received higher credibility ratings when they could clearly demonstrate they were accurate. In contrast, our behavioral measure of persuasion showed a stronger effect for confidence. Participants reported tentative credibility for confident advisors whom they could not verify as accurate but were still highly persuaded by them. This could be due to the lack of other advice or information they could use in making their final estimates, as well as differences in implicit and explicit cognitions.

When feedback was available at a cost, perceived advisor credibility was driven by accuracy and remained low for advisors of unknown accuracy. With respect to actual influence, although accuracy played an important role when feedback was free, participants relied more heavily on the confidence heuristic as it became more difficult and costly to assess the accuracy of advice (support for *Hypothesis 3*). We also found that participants were more inclined to scrutinize low confidence advice by buying feedback on its accuracy.

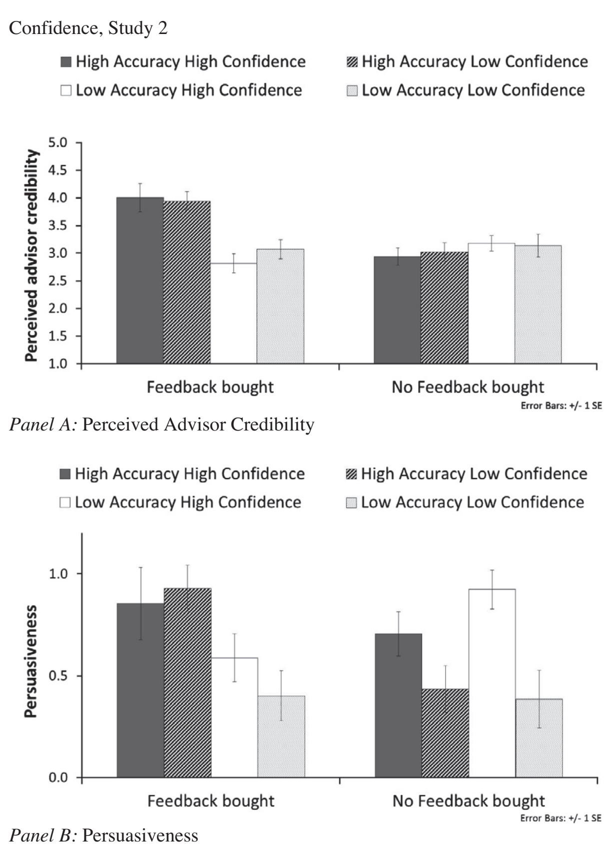


Fig. 4. Costly feedback condition: effect of buying feedback by advisor accuracy and confidence, study 2.

In some situations, we have ready access to the information we need to assess previous calibration, such as with the forecasts of meteorologists and sports handicappers. Consistent with our results, in such situations we would expect to see overconfidence backfire. But there are many other situations in which assessing a source's calibration is impossible, very difficult, or costly. It is often difficult to assess the predictions of political candidates and management consultants because they are too vague, have too many conditions, and the time to realization is too long. And in matters of opinion, there are rarely criteria clear enough that any piece of advice could be deemed unequivocally correct or incorrect (Kerr, MacCoun, & Kramer, 1996; Laughlin & Ellis, 1986; Van Swol, 2011). In such situations, "cheap talk" may indeed pay off for the advisor, particularly with "actual" influence, if not with perceived credibility.

General discussion

The current research contributes to the debate about the consequences of expressing confidence (Anderson et al., 2012; Price & Stone, 2004; Tenney et al., 2007, 2008). Our two studies help to clarify the role of source calibration and the boundaries on its effects by demonstrating that feedback on advisor performance is a key moderator variable.

If advice is easy to calibrate (as in Study 1 and the free feedback condition in Study 2), people do attend to accuracy and so it moderates the effect of confidence for perceived credibility. These results suggest some optimism regarding people's ability to call "bull" on advisors who indulge in cheap claims of confidence without delivering corresponding accuracy. However, this calibration seems specific to a situation in which feedback is clear. Evidence for the calibration hypothesis is also most evident with explicit advisor credibility ratings; even then, it is most reliably evident when focusing on the overconfident advisor. Similarly, in previous work on the calibration hypothesis (Tenney et al., 2007, 2008), the focus of the calibration was on the overconfident individual (who lost substantial credibility when found to be inaccurate) rather than the underconfident individual (similarly mis-calibrated by displaying low confidence when accurate).[7] Study 2's free feedback condition revealed that accurate advisors were not punished for poor calibration whereas high confidence inaccurate advisors were, suggesting that accuracy may be more important than calibration in low confidence advisors. The importance of accuracy was also seen in our behavioral measure of persuasiveness where it was the sole determinant of this measure of influence.

If accuracy information is not available, people use confidence as a cue or proxy for accuracy. They rate highly confident advisors as more credible and are more persuaded by them. In the absence of feedback, Price and Stone's (2004) confidence heuristic aligns with Tenney et al.'s (2008) "presumption of calibration" hypothesis; if we presume calibration (until we learn otherwise), then confidence does imply accuracy. However, since credibility ratings for the confident and accurate advisors were significantly higher when accuracy is known, this presumption can only be partial at best. Advisors do not get the full benefit of the doubt with perceived credibility, but they remain just as persuasive as noted in our

[7] Tenney and Spellman (2011) is an exception: In this paper, the authors found self-knowledge (i.e. good calibration) to be more beneficial than being modest. However, the "modest" individuals in these experiments also made high confidence errors (i.e., were confident when wrong and uncertain when right).

behavioral measure. The confidence heuristic may be more evident in actual influence or other implicit measures, in contrast to more explicit measures that show some caution in rating advisors as highly credible in the absence of feedback.

The results from our "costly feedback" condition in Study 2 are particularly interesting. Here, participants' behavior also tells a pessimistic story. First, when feedback was costly, participants were more persuaded by the confident advisor, even though they indicated that accuracy was more important in determining an advisor's credibility. Second, participants were actually less likely to seek verification of an advisor's accuracy when the advice was delivered with confidence.

Thus, feedback moderated the extent to which participants used the confidence heuristic. Furthermore, even when participants acknowledged that accuracy should drive credibility and were cautious in rewarding high credibility to high confidence advisors with unknown accuracy, they were still more persuaded by these confident advisors (seen in no-feedback and costly feedback conditions) and did not check on their accuracy as readily (seen in costly feedback condition). These studies shed some light on how our society rewards the expression of confidence, particularly if people do not bother to check on accuracy when they have the opportunity to do so (Anderson et al., 2012).

Our findings from the free feedback condition (which show support for calibration over confidence, for credibility ratings—particularly for high confidence advisors) lead to the question of why Price and Stone (2004) found evidence for the confidence heuristic with feedback and repeated trials. We offer several viable explanations.

First, the participants in Price and Stone's study did not have to make their own judgments on the direction of stock movement and did not appear to have real monetary incentives to be accurate regarding their preferred choice of advisor or their recollection of the advisor's performance. The participants could, therefore, be simply less attentive to the accuracy information. They could take a shortcut to assess advisors via the confidence heuristic since it was too effortful (or costly) for them to calibrate.

Second, while Price and Stone provided feedback about advisor performance, they did not systematically vary accuracy. Both the moderate and overconfident advisors in Price and Stone's paradigm gave the same direction for whether the stock would go up or down, and the overconfident advisor added an extra 15% confidence in that direction. For example, if the stock had a 65% chance of going up and the moderate advisor stated correctly '65%', the overconfident advisor would say '80%', and a resulting answer to the advisee could be that the stock went up. Out of 24 trials, the advisors were correct in 18 of them. Even if the moderate advisor is better-calibrated than the advisor who makes more extreme predictions according to normative probability calculus, it could appear to the participants examining each trial (rather than the average over 24 trials) that the overconfident advisor is "more accurate, more often" (Keren & Teigen, 2001). Price and Stone likewise suggest "an alternative interpretation [to the confidence heuristic] is that participants preferred the extreme advisor because they found his judgments easier to discriminate" (2004, p. 44). Because our weight measure was continuous (as opposed to dichotomous), our approach allows us to distinguish extreme confidence from discriminability with respect to accuracy, and we find strong evidence that accuracy is more important than confidence (and perhaps calibration) when performance data are clear and freely available.

Third, our study employed a large discrepancy in advisor confidence with high confidence advisors averaging at 95% confidence and low confidence advisors at 30%. The differences in confidence between the two advisors in Price and Stone's study were smaller and it is possible that the overconfident advisor was simply not overconfident enough for advisees to register poor calibration. This may also suggest that some overconfidence may be beneficial to advisors by generating greater influence as long as people are not registering (or calibrating) this overconfidence.

Limitations

Participants in our studies got advice in written form, with no other interaction between advisor and advisee. This limited communication allowed for the tight experimental control we needed in order to make strong causal conclusions about what aspect of confidence makes advice credible, without the complicating attributions that would come with the display of confidence through para-verbal or non-verbal behavior (Anderson et al., 2012; Anderson & Kilduff, 2009; Ridgeway & Correll, 2006; Ridgeway, 2001). Although we have not attempted to explore verbal interactions in this paper, we should note that, even with face-to-face interaction, Van Swol and Sniezek (2005) found that *written* expressions of confidence were the most significant factor affecting advisees.

Our study did not explore other factors that could affect advice-taking such as advisor incentives to give accurate or inaccurate advice. Knowing your advisor is paid to be accurate is likely to alter how much you trust your advisor (Sah, Loewenstein, & Cain, 2013). Finally, given cultural differences in the norms surrounding the expression and interpretation of confidence (Li & Fang, 2004; Lundeberg, Fox, Brown, & Elbedour, 2000), the effects of advisor confidence and accuracy may well vary in different cultures.

Increasing your influence

Every advisor and communicator, from writers of scholarly articles to presidents, must make choices about how much confidence to attach to their conclusions. The confidence their communication inspires in others has a direct influence on the number of citations or votes they get. And so the question posed in this article is of profound practical importance.

If behaving confidently were all that it took to gain others' confidence, the lesson for authors, advisors, and presidents would be simple: Express maximal confidence all the time. Indeed, "cheap talk" claims of confidence can increase one's influence over others when measures of accuracy are not available or it is costly to obtain them. However, if easy-to-calibrate performance data is free, clear, and available, accuracy appears to be an important component in determining credibility and influence. If you suspect your advice may turn out to be mistaken it is better to lower your confidence. We find that modest advisors do not suffer greatly if they are found to be accurate (or inaccurate), whereas errors made with great confidence have a large detrimental effect on credibility. Furthermore, offering accuracy data on your own performance increases your credibility if you can show you are correct.

Can we offer any advice to the consumers of advice? We should seek to reward our advisors for their accuracy rather than their confidence, something we can only do if we seek out accuracy information. This does not mean focusing only on results, because in stochastic environments we can't reasonably expect our advisors to get it right every time. But we can insist that they be properly calibrated in the confidence with which they offer their advice.

Acknowledgments

The authors wish to thank Jack Soll, Elizabeth Tenney and Lamar Reinsch for their helpful comments on an earlier version of this manuscript.

References

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*(5), 888–918.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.

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

Anderson, C., & Kilduff, G. J. (2009). Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *Journal of Personality and Social Psychology, 96*(2), 491–503.

Berman, G. L., & Cutler, B. L. (1996). Effects of inconsistencies in eyewitness testimony on mock-juror decision making. *Journal of Applied Psychology, 81*(2), 170–177.

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes, 101*(2), 127–151.

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26*(3), 353–364.

Brewer, N., & Wells, G. L. (2006). The confidence–accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied, 12*(1), 11–30.

Chaiken, S., & Eagly, A. H. (1976). Communication modality as a determinant of message persuasiveness and message comprehensibility. *Journal of Personality and Social Psychology, 34*(4), 605–614.

Deffenbacher, K. A. (1980). Eyewitness accuracy and confidence: Can we infer anything about their relationship? *Law and Human Behavior, 4*(4), 243–260.

Dillard, J. P., Weber, K. M., & Vail, R. G. (2007). The relationship between the perceived and actual effectiveness of persuasive messages: A meta-analysis with implications for formative campaign research. *Journal of Communication, 57*(4), 613–631.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*(9), 913–920.

Evans, J. S. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454–459.

Fischhoff, B., & Bruine De Bruin, W. (1999). Fifty–fifty = 50%? *Journal of Behavioral Decision Making, 12*(2), 149–163.

Gino, F., & Moore, D. A. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making, 20*(1), 21–35.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*(1), 17–41.

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

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697–720.

Kassin, S. M. (1985). Eyewitness identification: Retrospective self-awareness and the accuracy–confidence correlation. *Journal of Personality and Social Psychology, 49*(4), 878–893.

Keren, G., & Teigen, K. H. (2001). Why is p = .90 better than p = .70? Preference for definitive predictions by lay consumers of probability judgments. *Psychonomic Bulletin and Review, 8*(2), 191–202.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review, 103*(4), 687–719.

Klayman, J., Soll, J. B., González-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.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*(3), 177–189.

Li, S., & Fang, Y. (2004). Respondents in Asian cultures (eg, Chinese) are more risk-seeking and more overconfident than respondents in other cultures (eg, in United States) but the reciprocal predictions are in total opposition: How and why? *Journal of Cognition and Culture, 4*(2), 263–292.

Lundeberg, M. A., Fox, P. W., Brown, A. C., & Elbedour, S. (2000). Cultural influences on confidence: Country and gender. *Journal of Educational Psychology, 92*(1), 152–159.

Masicampo, E. J., & Baumeister, R. F. (2008). Toward a physiology of dual-process reasoning and judgment. *Psychological Science, 19*(3), 255–260.

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

Moore, D. A., & Klein, W. M. (2008). Use of absolute and comparative performance feedback in absolute and comparative judgments and decisions. *Organizational Behavior and Human Decision Processes, 107*(1), 60–74.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231–260.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1990). The adaptive decision maker: Effort and accuracy in choice. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn* (pp. 129–153). Chicago: University of Chicago Press.

Phillips, J. M. (1999). Antecedents of leader utilization of staff input in decision-making teams. *Organizational Behavior and Human Decision Processes, 77*(3), 215–242.

Price, P. C., & Stone, E. R. (2004). Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic. *Journal of Behavioral Decision Making, 17*(1), 39–57.

Ridgeway, C. L., & Correll, S. J. (2006). Consensus and the creation of status beliefs. *Social Forces, 85*(1), 431–453.

Ridgeway, C. L. (2001). Social status and group structure. In *Blackwell handbook of social psychology: Group processes* (pp. 352–375).

Sah, S., & Loewenstein, G. (2012). More affected = more neglected: Amplification of bias in advice to the unidentified and many. *Social Psychological and Personality Science, 3*(3), 365–372.

Sah, S., Loewenstein, G., & Cain, D. M. (2013). The burden of disclosure: Increased compliance with distrusted advice. *Journal of Personality and Social Psychology, 104*(2), 289–304. http://dx.doi.org/10.1037/a0030527.

Shaw, J. S., & McClure, K. A. (1996). Repeated postevent questioning can lead to elevated levels of eyewitness confidence. *Law and Human Behavior, 20*(6), 629–653.

Smith, J. F., Mitchell, T. R., & Beach, L. R. (1982). A cost-benefit mechanism for selecting problem-solving strategies: Some extensions and empirical tests. *Organizational Behavior and Human Performance, 29*(3), 370–396.

Sniezek, J. A. (1992). Groups under uncertainty: An examination of confidence in group decision making. *Organizational Behavior and Human Decision Processes, 52*(1), 124–155.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge-advisor decision making. *Organizational Behavior and Human Decision Processes, 62*(2), 159–174.

Sniezek, J. A., & Van Swol, L. M. (2001). Trust, confidence, and expertise in a judge-advisor system. *Organizational Behavior and Human Decision Processes, 84*(2), 288–307.

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., & Spellman, B. A. (2011). Complex social consequences of self-knowledge. *Social Psychological and Personality Science, 2*(4), 343–350.

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*(5), 1368–1375.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

Van Swol, L. M. (2009). The effects of confidence and advisor motives on advice utilization. *Communication Research, 36*(6), 857–873.

Van Swol, L. M. (2011). Forecasting another's enjoyment versus giving the right answer: Trust, shared values, task effects, and confidence in improving the acceptance of advice. *International Journal of Forecasting, 27*(1), 103–120.

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

Wells, G. L. (1993). What do we know about eyewitness identification? *American Psychologist, 48*(5), 553–571.

Wells, G. L., Lindsay, R. C., & Ferguson, T. J. (1979). Accuracy, confidence, and juror perceptions in eyewitness identification. *Journal of Applied Psychology, 64*(4), 440–448.

Yaniv, I. (1997). Weighting and trimming: Heuristics for aggregating judgments under uncertainty. *Organizational Behavior and Human Decision Processes, 69*(3), 237–249.

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

Yates, J. F., Price, P. C., Lee, J. W., & Ramirez, J. (1996). Good probabilistic forecasters: The ''consumer''s' perspective. *International Journal of Forecasting, 12*(1), 41–56.

Zarnoth, P., & Sniezek, J. A. (1997). The social influence of confidence in group decision making. *Journal of Experimental Social Psychology, 33*(4), 345–366.