AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 1, 159–170
https://doi.org/10.1037/xge0001487

Premature Predictions: Accurate Forecasters Are Not Viewed as More Competent for Earlier Predictions

Robert Mislavsky[1] and Celia Gaertig[2]
[1] Marketing Area, Carey Business School, Johns Hopkins University
[2] Marketing Group, Haas School of Business, University of California, Berkeley

How does the timing of a prediction influence how a forecaster is perceived? Many people believe that they will be seen as more competent if they make accurate predictions far in advance of an event. However, we find that forecasters are not seen as more competent—and are sometimes seen as less competent—when they make predictions far in advance of an event occurring. Furthermore, we find that this is because observers recognize that events far in the future are less knowable, suggesting that they may attribute accurate but premature forecasts more to luck than to skill. Forecasters would benefit from knowing this when considering making predictions. They are not penalized for waiting until more information is known but may lose credibility if they make a prediction too early.

Public Significance Statement
Forecasting has become increasingly prevalent and increasingly lucrative for successful forecasters. However, there has been little research on how people form judgments of forecasters themselves. This article examines how one component of a forecast, when the forecast is made, influences a forecaster's perceived competence. Specifically, we find that, although making earlier predictions is more difficult, forecasters are not perceived as more competent the earlier they make predictions (and may be seen as less competent if they make predictions too early). Our findings suggest that forecasters (whether professional or amateur) may be better off waiting to make predictions until significant information is available to make a truly informed prediction.

Keywords: uncertainty, forecasting, predictions, open science

Supplemental materials: https://doi.org/10.1037/xge0001487.supp

Forecasts play an important role in personal and societal decision making. We regularly listen to forecasters provide predictions about the weather, stock prices, sporting events, political elections, and the trajectory of a pandemic, among other future events. These forecasts are typically made with an audience in mind, and there is good reason to believe that forecasters care about how they and their forecasts are perceived. But what influences perceptions of forecasters?

Prior research investigating social perceptions of forecasters has primarily focused on the content of the prediction itself and how it is delivered. For example, people value forecasters who are accurate and confident (e.g., Price & Stone, 2004; Radzevick & Moore, 2011; Tenney et al., 2008). However, many more factors may influence observers' judgments of a forecaster. Importantly, because predictions are necessarily directed at events that occur in the future, different forecasters may make predictions at different points in time prior to the event. For example, a forecaster may make a prediction about who will win next year's Super Bowl right now or closer to when the game takes place. This raises the question of whether the timing of a forecaster's prediction affects how they are evaluated. Because less information about the likely outcome is known the farther in advance a prediction is made, an accurate prediction becomes ostensibly more difficult. In this case, it is reasonable to believe that forecasters may benefit reputationally from accurately making earlier (and therefore more difficult) predictions.

This article was published Online First October 23, 2023.
Robert Mislavsky https://orcid.org/0000-0002-9620-3528
This article was previously presented at the 2020 Society for Judgment and Decision Making Conference and the 2021 Association for Consumer Research Conference. The authors thank Johns Hopkins Carey Business School, the Haas School of Business at the University of California, Berkeley, and the Beatrice Foods Co. Faculty Research Fund at the University of Chicago Booth School of Business for financial support. Data, analysis code, materials, and preregistrations are available at https://researchbox.org/354.

Robert Mislavsky served as lead for conceptualization, formal analysis, and methodology. Celia Gaertig served in a supporting role for conceptualization, formal analysis, and methodology. Robert Mislavsky and Celia Gaertig contributed equally to writing–original draft and writing–review and editing.

Correspondence concerning this article should be addressed to Robert Mislavsky, Carey Business School, Johns Hopkins University, 100 International Drive, Baltimore, MD 21202, United States. Email: mislavsky@jhu.edu

However, some predictions seem to be provided prematurely. To illustrate, consider the ubiquitous example of election predictions.[1] On November 3, 2020, Joe Biden was declared the winner of the United States presidential election. Within days, several news articles and opinion pieces were published speculating on who might win the 2024 presidential race. This was a full 4 years before the next election would take place and 2 years before the first major candidate declared his intention to run[2] (Orr et al., 2022; Peek, 2020; Z. Jones, 2020). Such premature predictions are not limited to political horseraces. Indeed, major publications regularly make predictions well before having adequate information in domains such as entertainment (Davis, 2021; McIntyre, 2021), sports (Borzello, 2021), and investing (Bambrough, 2020; Lefton, 2018). Even if these forecasters did not themselves decide to make predictions so early, but were instead asked to do so, understanding how the timing of a prediction affects observers' judgments of the forecaster can offer important insights into an aspect of forecaster perception that is currently not well understood.

In the current research, we examine the social and reputational consequences of making (accurate) predictions at different points in time before an event takes place. We show that contrary to people's lay intuitions, forecasters do not receive more credit for making predictions far in advance of an event. We discuss the role of epistemicness (i.e., how knowable an outcome is at the time of the prediction) in this context and show that, when evaluating others' forecasts, observers recognize that events far in the future are less knowable than events closer in time, and that this mediates the effect of prediction timing on perceived competence. Similarly, for events that are considered knowable far in advance, we do not find any effect of prediction timing on a forecaster's perceived competence.

What Do Observers Value in a Forecaster?

Although predictions may be made in private (such as an individual investor predicting which of two stocks will provide a greater financial return), they are often made to an audience. Expert consultants provide forecasts to their clients, political pundits offer predictions about who will win an upcoming election to the public, and friends debate who will win the Super Bowl in the upcoming season. As such, it is important to understand how observers perceive forecasts and how this may impact their judgments of the forecaster.

A large amount of research has shown that even qualitatively identical forecasts may be perceived differently depending on how they are presented to the audience. For example, the same forecast may be perceived differently depending on whether it is framed as a numerical (e.g., 60% chance) or a verbal probability (e.g., likely; Beyth-Marom, 1982; Dhami & Mandel, 2022; Lichtenstein & Newman, 1967; Windschitl & Wells, 1996), and this has consequences for how people incorporate forecasts into their own judgments (Mislavsky & Gaertig, 2022). Furthermore, simply changing the context in which a forecast is presented is sufficient to change people's understanding of it (Windschitl & Weber, 1999).

How a forecast is presented also influences observers' judgments of the forecasters. For example, observers perceive forecasters who use higher rather than lower numbers as more confident and more trustworthy (Bagchi & Ince, 2016). In general, people seem to value confidence in forecasters (e.g., Price & Stone, 2004; Radzevick & Moore, 2011), and forecasters seem to have internalized this, often providing estimates that are too certain (Anderson et al., 2012; Van Zant, 2022). However, overconfident forecasts are not always well perceived. When feedback is available and observers can judge the accuracy of a forecast, they tend to value calibration more than confidence (Tenney et al., 2008). Overconfidence can also be a liability in inherently uncertain environments, in which observers prefer forecasts that are appropriately uncertain to overly certain forecasts (Gaertig & Simmons, 2018, 2023; Gustafson & Rice, 2019; Howe et al., 2019). This is also consistent with the finding that people prefer judgments that balance both informativeness and accuracy (i.e., having a wide enough range to contain the actual outcome; Yaniv & Foster, 1995).

Taken together, prior work investigating social judgments of forecasters primarily focused on the content of a forecast and how it is presented. In the current research, we examine how the context in which a prediction is made (specifically, when it is predicted) influences perceptions of the forecasters.

Prediction Timing and the Epistemicness of Future Outcomes

In our research, we investigate whether prediction timing (i.e., when a prediction is made) influences observers' judgments of a forecaster's competence and credibility. While forecasters may not always have control over the timing of a prediction (e.g., they may be asked to make a prediction), the timing of a prediction may have important consequences for how forecasters are evaluated. Making a prediction about the outcome of an event far in advance of that event occurring is ostensibly more difficult, as there is less information available to the forecaster. Because of this, it is possible that observers will view forecasters more favorably for making an accurate prediction far in advance of an event.

At the same time, the outcomes of events in the far future are less knowable. That is, adding more time between a prediction and an event can make an event's outcome less epistemic (i.e., less knowable in advance) and more aleatory (i.e., more influenced by randomness; Tannenbaum et al., 2017). Observers who intuit this may attribute correct, but premature predictions to luck rather than the forecaster's ability. This, in turn, could mean that observers will give forecasters less credit for making a correct prediction far in advance of an event. If this is the case, forecasters trying to maximize perceptions of competence would not benefit from making a prediction as early as possible. Instead, they may be better off waiting to make a forecast until more relevant information pertaining to the prediction is available.

Overview of Studies

Our research investigates whether accurate forecasters are perceived as more competent the earlier they make their predictions. We find that, despite their intuitions, they are not, and may even be perceived as less competent when they make extremely early predictions.

Across six studies, we asked participants either to predict how they would be judged for having made a (correct) prediction at different points in time (Studies 1a and 1b) or to judge a forecaster who made a prediction (Studies 2–5). In all studies, we used stimulus

[1] Watts and Rothschild (2017) found that "horserace" stories (i.e., stories that dealt with the state of the campaign and candidates' likelihoods of winning) accounted for over half the New York Times' front-page election coverage during the 2016 U.S. Presidential election.

[2] Donald Trump launched his campaign on November 15, 2022.

sampling to include different predictions in order to ensure that our effects were not driven by an idiosyncratic feature of any specific prediction (Wells & Windschitl, 1999). As preregistered, our analyses for each study are presented collapsed across stimuli. We provide analyses by individual stimuli in Supplement 2 in the online supplemental materials.

In Studies 1a and 1b, participants report that they expect to receive more credit (Study 1a) and to be perceived as more competent (Study 1b) the farther in advance they make a correct prediction. However, in Studies 2 and 3, participants give forecasters less credit on average for earlier correct predictions, although this is primarily driven by extremely early forecasts. We find that this is due to observers perceiving that events are less knowable far in the future. Importantly, Study 4 shows that this effect also holds for events where the outcome is still uncertain, and Study 5 shows that the effect holds when predictions are prompted by others, rather than volunteered by the expert.

Transparency and Openness

We preregistered all studies. For all studies, we report all data exclusions, all manipulations, and all measures. Sample sizes were determined before data collection. Supplemental materials, including data, analysis code, preregistrations, and survey materials, are available at https://researchbox.org/354.

Studies 1a and 1b—Participants Expect More Credit for Earlier Predictions

Studies 1a and 1b investigate participants' expectations of how much credit they would receive (i.e., how competent others would think they are) for correct predictions made at different times before the event.

Study 1a—Within-Subjects Expectations

Sample

We recruited 224 participants on Amazon's Mechanical Turk (MTurk), of which 204 (52.7% male, 46.8% female, 0.0% another option not listed; $M_{age} = 36.7$ years)[3] passed an attention check embedded at the beginning of the study. Participants were paid $0.40 to complete the survey.

Design

In this study, we manipulated how far in advance participants imagined having made a prediction within subjects. Specifically, participants imagined having made a correct prediction either 10 years, 5 years, 1 year, or 1 month in advance.

Participants were asked to imagine that they were making a prediction about an unspecified future event. They were provided with examples of such a prediction (e.g., a prediction about who will win a certain election, whether a stock's price will be above a certain amount, or whether certain countries will go to war) and were told to assume that whatever prediction they made would be correct. We then asked participants to indicate how much credit they thought they would get from others if they made this correct prediction at different times before the event. Specifically, we asked, "How much credit do you think you would get from others (e.g., how smart will others think you are) if you made this prediction (10 years/5 years/1 year/1 month) in advance and it was correct?" (1 = *no credit at all*; 7 = *a lot of credit*). All participants answered this question for four different time frames (10 years, 5 years, 1 year, and 1 month in advance), with the time frames presented in descending order (i.e., 10 years first, 1 month last).

Results

We used ordinary least squares (OLS) regression to regress expected credit on how far in advance the prediction was made (in months; e.g., 1 year = 12 months), clustering at the participant level.

Participants expected to receive more credit for a correct prediction the earlier the prediction was made. Specifically, the amount of expected credit increased from 4.47 out of 7 ($SD = 1.84$) when the prediction was made 1 month in advance to 5.09 ($SD = 1.86$) when the prediction was made 10 years in advance, $b = 0.005$, $t(201) = 3.06$, $p = .002$.

Because this study used a within-subjects manipulation of prediction timing, we can also measure the proportion of participants who have the expectation that they will receive more credit for earlier predictions (and the proportion who show other patterns).[4] Exactly half of the participants (50.0%) expected to receive more credit if they made increasingly early predictions. Slightly more than one quarter of participants (27.7%) expected to receive less credit for earlier predictions, and 10.9% expected the same amount of credit regardless of the prediction timing (compared to increasing expectations: $Zs > 4.59$, $ps < .001$). The remaining participants (11.4%) showed no particular pattern.

While Study 1a provides evidence that people generally expect more credit for earlier predictions, the within-subjects design may have created a demand effect where participants felt that they should report expecting more credit for earlier predictions. Additionally, since we did not tell participants to imagine a specific prediction, they may have imagined making a prediction about an outcome that would have been reasonably knowable 10 years in advance. Study 1b addresses these concerns by manipulating prediction timing between subjects and providing participants with specific predictions.

Study 1b—Between-Subjects Expectations

Sample

We recruited 1,311 participants on Prolific, of which 917[5] (51.2% female, 47.9% male, 0.8% another option not listed; $M_{age} = 38.4$

[3] Participants were asked to indicate their gender in a multiple-choice question at the end of the survey, where the options were "male," "female," "another option not listed," and "prefer not to say." They were asked to indicate their age by typing in a number in an open-ended question. Answering these questions was optional, and we did not ask our participants any other demographic questions.

[4] This analysis was suggested during a lab meeting and was therefore not preregistered. For this analysis, we defined "more credit" as participants choosing a greater number for 10 years than for 1 month and the same or greater number for 10 years than for 5 years, for 5 years than for 1 year, and so on. We defined "less credit" similarly in the opposite direction.

[5] We preregistered that we would collect data from 800 participants. However, an error in the survey caused some participants to not receive the correct Prolific approval link. As a result, more participants were able to complete the survey than we preregistered.

years) passed an attention check that was embedded at the beginning of the study. Participants were paid $0.70 to complete the survey.

Design

Participants were assigned to one of 24 conditions in a 6 (domain: politics vs. sports vs. investments vs. Oscars vs. Brexit vs. business) × 4 (prediction timing: 1 month vs. 1 year vs. 5 years vs. 10 years) between-subjects design.

Participants were told to imagine that they made a (correct) prediction in one of six domains:

1. *Politics*—Joe Biden will be elected President of the United States in November 2020.
2. *Sports*—The Tampa Bay Buccaneers will win Super Bowl 55 in February 2021.
3. *Investments*—The price of Bitcoin will be above $30,000 in January 2021.
4. *Oscars*—A non-English language film will win Best Picture at the Academy Awards in 2020.
5. *Brexit*—The United Kingdom will formally withdraw from the European Union in January 2020.
6. *Business*—Jeff Bezos will step down from his role as Amazon CEO in 2021.

Participants then read that they made that prediction either 1 month, 1 year, 5 years, or 10 years before the event actually occurred, which we operationalized by manipulating the month and year that the prediction was made in the scenario text. For example, a prediction about the 2020 Oscars made 5 years in advance of the event read (emphasis added; full stimuli for all conditions available at https://researchbox.org/354):

> Imagine that you made a prediction about an event.
>
> Specifically, you predicted that **a non-English language film would win Best Picture at the Academy Awards in 2020**, which was correct.
>
> You made this prediction in **February 2015**, which was **5 years** before it actually happened.

To measure perceived credit, participants then answered four questions (adapted from Gaertig & Simmons, 2018) about how others would perceive them. Specifically, they indicated how knowledgeable, credible, and competent others would perceive them and how much others would trust them, all on 1 (*not at all*) to 7 (*extremely*) scales. For example, participants in the example shown above answered the following questions:

1. How knowledgeable would people think you are now if you had made this prediction in **February 2015** (**5 years** in advance)?
2. How credible would people think you are now if you had made this prediction in **February 2015** (**5 years** in advance)?
3. How competent would people think you are now if you had made this prediction in **February 2015** (**5 years** in advance)?
4. How much do you think people would trust you now if you had made this prediction in **February 2015** (**5 years** in advance)?

The bolded parts in the questions above were changed according to the assigned prediction timing condition (i.e., 1 month, 1 year, 5 years, or 10 years). We averaged the four dependent variables to create an index of expected credit ($\alpha = .93$).

Results

As preregistered, we collapsed results across the six prediction domains. We used OLS regression to regress expected credit on months, using fixed effects for prediction domain.

As in Study 1a, participants expected more credit for a correct prediction the earlier the prediction was made. The amount of expected credit increased from 4.88 out of 7 ($SD = 1.03$) when the prediction was made 1 month in advance to 5.22 ($SD = 1.41$) when the prediction was made 10 years in advance, $b = 0.002$, $t(903) = 2.71$, $p = .007$ (Figure 1). Although we did not power this study to detect an effect for each domain, this pattern held directionally for four out of the six domains and significantly for three out of the six domains (politics, investments, and Brexit; no domains were significant in the opposite direction, see Supplement S2 in the online supplemental materials).

Study 2—Forecasters Get Less Credit for Very Early Predictions

In Studies 1a and 1b, we show that participants themselves expect to receive more credit for their own (accurate) early predictions. But are they correct? We test this question in Study 2 and find that forecasters are not seen as more competent when making earlier predictions and are perceived as less competent when their forecasts are made very far in advance. Additionally, we find that this effect is driven by observers perceiving that events far in the future are less knowable.

Sample

We recruited 1,389 participants on Prolific, of which 992 (52.0% male, 46.1% female, 1.0% another option not listed; $M_{age} = 35.5$ years) passed an attention check that was embedded at the beginning of the study. Participants were paid $0.70 to complete the survey.

Design

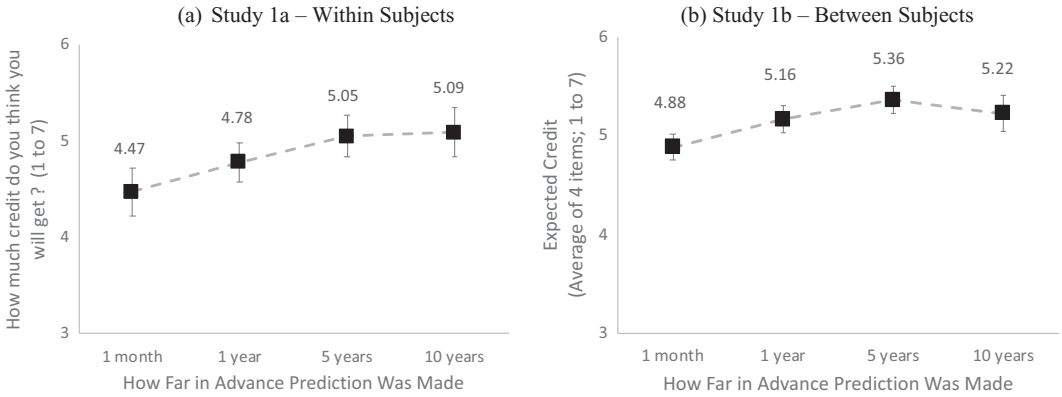
We randomly assigned participants to one of 20 conditions in a 5 (domain: politics vs. sports vs. business vs. Oscars vs. Brexit) × 4 (prediction timing: 1 month vs. 1 year vs. 5 years vs. 10 years) between-subjects design.

All participants in this study judged an expert who made a (correct) prediction about a specific event. Participants were asked to imagine that they were reading an article where an expert had offered that prediction. We manipulated both the prediction domain and how far in advance of the event the prediction was made. The prediction domains were the same as in Study 1b, except without the investments prediction.[6] We manipulated prediction timing by telling participants that the expert's prediction was made either 1 month, 1 year, 5 years, or 10 years before the event happened. In all conditions, participants were told that the expert's prediction was correct.

[6] This study was run chronologically before Study 1b. The investments prediction was added to Study 1b to improve stimulus sampling.

Figure 1

Participants Expect More Credit for Earlier Predictions



Note. Results from Studies 1a (a) and 1b (b). Error bars represent 95% confidence intervals.

For example, participants assigned to the "sports/12 months in advance" condition read the following (emphasis added; full stimuli for all conditions available at https://researchbox.org/354):

> Imagine that you are reading an article published in **February 2020**. An expert is quoted in the article as saying:
>
> **"The Tampa Bay Buccaneers will win Super Bowl 55 in February 2021."**
>
> The prediction was correct. The Tampa Bay Buccaneers won Super Bowl 55 on February 7, 2021.

After seeing the prediction and the outcome, participants answered five questions evaluating the competence of the expert, which we combined into a single index of forecaster competence (adapted from Gaertig & Simmons, 2018; see Table 1; $\alpha = .92$).

Additionally, we measured participants' perceptions of how knowable the event was at the time the prediction was made by asking participants four questions adapted from the Epistemic-Aleatory Rating Scale (EARS; Tannenbaum et al., 2017). The EARS asks participants to indicate whether they perceive an event as being more epistemic (knowable in advance) or aleatory (influenced by chance factors), with higher numbers indicating that participants feel the outcome of the event is more knowable in advance and less influenced by randomness. We selected four items from the EARS and slightly modified their wording to highlight that we are asking participants to rate the epistemicness of the event at the time the prediction was made. For example, one of the items read, "At the time that this prediction was made, this outcome was knowable in advance, given enough information" (1 = *not at all*; 7 = *very much*). Table 1 presents the full list of items that we used in this study. As per our preregistered analysis, we combined the four items into a single measure of perceived epistemicness ($\alpha = .77$).

We presented the dependent variables and EARS items to participants on separate pages in counterbalanced order (i.e., half of the participants answered the dependent variables [DVs] first and half answered the EARS first).

Results

Analysis Plan

As preregistered, we collapsed both the main dependent variable and epistemicness mediator across prediction domains and used OLS regression to regress each of the measures on how many months in advance the prediction was made, with fixed effects for prediction domain.

Main Analysis

Figure 2 shows the main results from Study 2. In contrast to the intuitions of participants in Studies 1a and 1b, Study 2 participants

Table 1

Measures Used in Studies 2–5

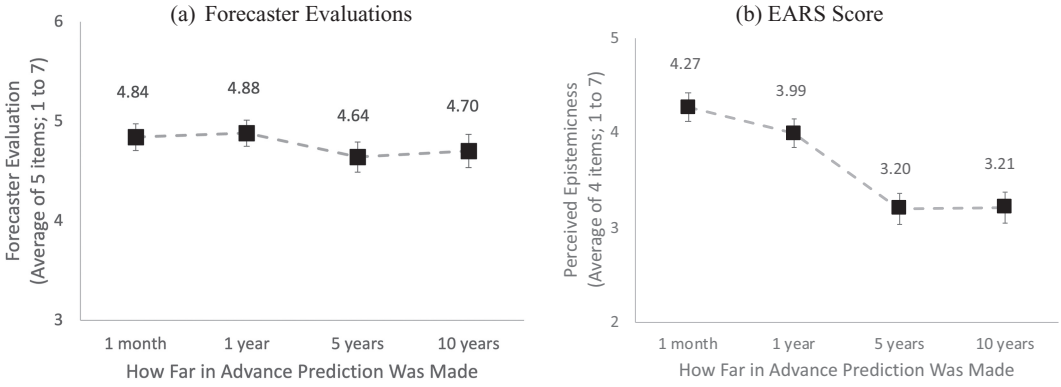
Dependent measure: Forecaster Evaluation (αs ≥ .92)
1. How knowledgeable is this person? (1 = *not at all*; 7 = *extremely*)
2. How credible is this person? (1 = *not at all*; 7 = *extremely*)
3. How competent is this person? (1 = *not at all*; 7 = *extremely*)
4. How much do you trust this person? (1 = *not at all*; 7 = *extremely*)
5. Would you seek additional information or advice from this person? (1 = *definitely not*; 7 = *definitely*)
Mediator: Adapted EARS (αs > .71)[a]
1. At the time that this prediction was made, this outcome was something that had an element of randomness. (1 = *not at all*; 7 = *very much; reverse-coded*)
2. At the time that this prediction was made, this outcome felt like it was determined by chance factors. (1 = *not at all*; 7 = *very much; reverse-coded*)
3. At the time that this prediction was made, this outcome was knowable in advance, given enough information. (1 = *not at all*; 7 = *very much*)
4. At the time that this prediction was made, this outcome was something that well-informed people would agree on. (1 = *not at all*; 7 = *very much*)

Note. EARS = Epistemic-Aleatory Rating Scale.

[a] In Study 4, we did not include the "at the time this prediction was made" qualifier in the EARS questions since the predictions were ostensibly made at the time of the study.

Figure 2

Forecasters Get Less Credit for Earlier Predictions (a), Which Are Perceived as Less Knowable (b)



Note. Study 2 results. (a) Participants' average evaluation of forecasters by how far in advance of the event the prediction was made. Higher numbers indicate a more positive evaluation. (b) The perceived epistemicness of the event. Higher numbers indicate that the event is more knowable in advance (vs. more influenced by chance factors). Perceived epistemicness mediates the effect of prediction timing on forecaster evaluation. Error bars represent 95% confidence intervals. EARS = Epistemic-Aleatory Rating Scale.

evaluated forecasters less positively, on average, the earlier the prediction was made. Specifically, participants' average evaluations decreased from 4.84 out of 7 ($SD = 1.08$) when the prediction was made 1 month in advance to 4.70 ($SD = 1.13$) when the prediction was made 10 years in advance, $b = -0.002$, $t(981) = -2.11$, $p = .035$. We present the results by stimulus in Supplement 2 in the online supplemental materials.

Mediation Analysis

Participants gave events lower EARS scores (i.e., rated them as less knowable) the more temporally distant they were from the experts' predictions. The average EARS score of the event decreased from 4.27 out of 7 ($SD = 1.20$) when the prediction was made 1 month in advance to 3.21 ($SD = 1.28$) when it was made 10 years in advance, $b = -0.009$, $t(981) = -10.63$, $p < .001$. When participants' responses to the EARS are included as a predictor of participants' evaluation of the forecaster, EARS mediates the main effect of prediction timing. For this study and all subsequent studies, we conducted any mediation analysis using 5,000 bootstrapped simulations with the PROCESS function (Hayes, 2018; Model 4 unless otherwise noted) from the "bruceR" R library (Bao, 2023). Here, the indirect effect of the event's EARS score excludes zero, 95% confidence interval (CI) [−0.002, −0.001].

While we preregistered that we would test mediation by using the EARS as one combined measure, we can also test whether the epistemicness items (Items 3 and 4, see Table 1) and aleatoriness items (Items 1 and 2, see Table 1) of the EARS have different mediating effects.[7] In this analysis, we find a significant indirect effect of the epistemicness measures, 95% CI [−0.0031, −0.0017] and a significant (but smaller) indirect effect of aleatoriness, [0.0002, 0.0012]. This indicates that the mediating effect of EARS seems to be driven primarily by participants believing that events far into the future are less knowable, rather than more likely to be affected by chance factors. Additionally, in subsequent studies, the indirect effect of epistemicness is consistently significant, but the indirect effect of aleatoriness is only significant in this study, supporting the idea that this is primarily an effect of perceived knowability.

Taken together, the results from Study 2 demonstrate that, contrary to people's intuitions, observers perceive forecasters who make predictions far in advance of an event as less competent. We furthermore show that this effect is mediated by observers perceiving events far into the future as less knowable. These results are also consistent with two earlier studies that we had run with different stimuli, and that we present in Studies S1 and S2 in Supplement 3 in the online supplemental materials.

Study 3—Events That Are Knowable Far in Advance

Study 2 shows that, in contrast to people's expectations, accurate forecasters are seen as less competent if they make forecasts far in advance. These diminished social perceptions result from observers acknowledging that the accuracy of such premature predictions may have been driven less by knowledge and more by luck.

It is possible that people consider the prediction events that we included in Study 2 to be events for which the relevant information needed to make an accurate prediction can only be known relatively close to the event occurring. For example, Study 2 included a prediction about the Super Bowl. In the months before the Super Bowl, much of the relevant information required to predict the winner has been realized. Schedules are set, and rosters and coaching staffs are generally known. However, accurately predicting the winner of the Super Bowl in *5 years* requires a much more complex series of events to go the forecaster's way. The predicted team may be given an atypically difficult schedule or may have its star player unexpectedly retire or be traded. As a result, successfully predicting events like these may become exponentially more difficult as the relevant timespan between the event and the forecast increases.

On the other hand, some predictions do not suffer from this problem to the same degree. Specifically, this is the case for forecasts for which

[7] We thank an anonymous reviewer for this suggestion.

short-term fluctuations are less likely to change the overall trend of an outcome. For example, an extremely hot or cold day (or even a large one-time increase in carbon emissions) is unlikely to change long-term global temperature trends. For these types of forecasts, participants should not perceive the epistemicness of the event to increase as dramatically as the timespan of the forecast increases. Therefore, if the perceived epistemicness of an event is indeed driving participants' evaluations of forecasters, the effect of prediction timing on forecaster evaluation should be reduced or eliminated for such prediction events that are more easily knowable far in advance (subsequently referred to as "long-term" predictions). In Study 3, we test this prediction. We first pretest different events to find out if participants perceive them as knowable far in advance. We then test whether having a prediction be knowable far in advance influences how much credit forecasters receive for an accurate prediction.

Pretest

In Study 3, our key manipulation is whether or not an event is "knowable" far in advance of an event. Since we did not want to rely on our own beliefs about whether events can be knowable far in advance, we conducted a two-stage pretest to develop the prediction stimuli for Study 3. In the first stage of this pretest (i.e., Pretest 1), we asked 175 participants on MTurk to provide us with some examples of predictions that seem "reasonable" to make at different points in time before the event (i.e., 1 day, 1 week, 1 month, 6 months, 1 year, 5 years, or 10 years in advance). From participants' responses in stage one (see data file on https://researchbox.org/354), we developed three predictions that were consistently considered "short-term" (i.e., predictable less than 1 week in advance) and three predictions that were consistently considered "long-term" (i.e., predictable more than 1 year in advance).

We next presented these stimuli to 109 participants in a separate study (Pretest 2) and asked them to indicate the earliest that they thought someone would have enough information to confidently make each prediction. We asked participants to select their answer from a multiple-choice list of 1 day before, 1 week before, 1 month before, 6 months before, 1 year before, or 5 years before. The purpose of this second stage was to confirm that the stimuli that we derived from the first stage were also considered "short term" and "long term" for a new set of participants. A majority of participants in Pretest 2 indicated that the predictions we intended to use as "short-term" predictions would be knowable 1 week or less in advance, while the "long-term" predictions would be knowable 1 year or more in advance. We then used these predictions as stimuli in Study 3, and we describe them below. We present the full details of the methods and results of the two pilot studies in Supplement 1 in the online supplemental materials.

Sample

For Study 3, we recruited 2,061 participants on MTurk, of which 1,740 (53.3% female, 45.7% male, 0.5% another option not listed; $M_{age} = 41.9$ years) passed an attention check that was embedded at the beginning of the study. Participants were paid $0.40 to complete the survey.

Design

The general design of this study was similar to the design of Study 2. Participants were randomly assigned to one of four conditions in a 2 (prediction type: knowable short-term vs. knowable long-term) × 2 (prediction timing: 1 week vs. 1 year) between-subjects design.

Just like in Study 2, participants judged an expert who made a (correct) prediction about a specific event happening in the future. However, in this study, we only included two prediction time frames: we manipulated prediction timing to be either 1 week or 1 year before the event. Importantly, we also manipulated whether or not the outcome predicted was reasonably knowable over the longer, 1 year time frame. Based on the results from our two-stage pilot study, we selected three short-term predictions (i.e., reasonably knowable 1 week in advance) and three long-term predictions (i.e., reasonably knowable a year or more in advance) for this study. The three short-term predictions (i.e., reasonably knowable 1 week in advance) were:

> *Temperature*—The high temperature in Washington, DC on October 16, 2020 will be between 65 and 70 degrees Fahrenheit.
>
> *Investments*—The NASDAQ will close between 11,500 and 12,000 on November 20, 2020.
>
> *Sports*—The Denver Broncos will beat the Miami Dolphins on November 22, 2020.

The three long-term predictions (i.e., reasonably knowable a year or more in advance) were:

> *Temperature*—The Earth's average temperature will be between 60 and 62 degrees Fahrenheit in 2019.
>
> *Economics*—The median household income in the United States (not adjusted for inflation) will pass $60,000 in 2018.
>
> *Population*—The population of Bangkok, Thailand will grow to over 10 million people in 2018.

Participants in all conditions learned that the prediction was correct and answered the same five items about the expert's perceived competence ($\alpha = .94$). They also answered the same four EARS items ($\alpha = .82$) that we used in Study 2. The dependent measure and EARS were presented in counterbalanced order.

At the end of the study, we included a manipulation check question by asking participants to answer the following question about the prediction timing: "In the scenario that you read on the previous page, how long before the event occurred did the expert make their prediction?" (answer choices: 1 week before vs. 1 year before). We preregistered that our analysis would only include participants who answered this question correctly. However, a substantial proportion of participants in the condition in which we presented them with a long-term prediction and told them the forecaster made the prediction on a date that was 1 week in advance answered the manipulation check question incorrectly (84.8%). Considering our stimuli again, we realized that the wording of this manipulation check question may not have been as clear for this condition (long-term prediction, 1 week in advance) compared to the other conditions. This is because, in contrast to the short-term predictions, the three long-term predictions did not specify the exact date at which the event occurred, but instead simply mentioned the year (e.g., "The Earth's average temperature will be between 60 and 62 degrees Fahrenheit in 2019"). At the same time, the date that we gave to participants in the 1-week-in-advance condition was always set to the end of the prior year (e.g., December 24, 2018). Hence, when deciding how to answer the question of whether

the prediction was made 1 week or 1 year in advance, participants may have felt ambiguity over the exact time at which the event in this condition actually occurred (i.e., whether they believe they should count the event as happening on the first day of the year or the last day of the year) and may have decided to go with the year difference (2018 vs. 2019) rather than the exact date, leading them to choose 1 year as the right answer.[8]

With this in mind and to provide the most conservative estimate of our effect, for our main analysis below, we present the results for all participants without excluding participants with incorrect manipulation check responses. However, our predicted effects are substantially larger when including only those who answered the manipulation check question correctly (see Supplement 2 in the online supplemental materials).

Results

Analysis Plan

As preregistered, we collapsed the three short-term events and three long-term events for our analysis. We used OLS regression to regress the respective measure (forecaster evaluation or EARS) on (a) whether the prediction was reasonably knowable in the short term or in the long term, (b) prediction timing (1 week vs. 1 year), and (c) the interaction of the two. The analysis included fixed effects for prediction domain.

Main Analysis

Participants in the reasonable short-term predictions condition rated forecasters less positively if they made predictions 1 year in advance ($M = 4.92$, $SD = 1.20$) compared to 1 week in advance ($M = 5.13$, $SD = 1.10$), $b = -0.207$, $t(835) = -2.71$, $p = .007$. This is in line with our previous results. However, when presented with reasonable long-term predictions, participants rated forecasters approximately equally if they made predictions 1 year in advance ($M = 5.35$, $SD = .99$) compared to 1 week in advance ($M = 5.32$, $SD = 1.00$), $b = 0.034$, $t(837) = .50$, $p = .619$. The interaction between event type and prediction timing was significant, $b = -0.241$, $t(1{,}672) = -2.36$, $p = .019$. That is, consistent with the idea that observers would consider how epistemic an event is in the long term, forecasters making predictions far in advance of an event seem to only face reputational consequences when predicting events that are not reasonably knowable at the time. They do not face the same negative judgment when predicting events that are considered more epistemic even in the long term. That is, forecasters seem to only face reputational consequences when they predict prematurely.

Mediation Analysis

In line with our predictions, participants in this study also felt that reasonable short-term prediction events were more knowable 1 week in advance ($M = 4.08$, $SD = 1.16$) compared to 1 year in advance ($M = 3.76$, $SD = 1.35$), $b = -0.322$, $t(830) = -3.99$, $p < .001$. However, for the reasonable long-term prediction events, there was no difference in the perceived epistemicness of these events being predicted 1 week ($M = 4.90$, $SD = 1.13$) versus 1 year ($M = 4.94$, $SD = 1.07$) in advance, $b = 0.041$, $t(835) = .55$, $p = .584$. The interaction between event type and prediction timing on EARS was significant, $b = -0.363$, $t(1{,}665) = -3.29$, $p = .001$.

When we include the EARS in a mediated moderation model (PROCESS Model 8), there is a significant indirect effect of the EARS score on short-term events, 95% CI [−0.110, −0.032], but not on long-term events, [−0.022, 0.040]. With both components of the EARS as separate mediators, we find a conditional indirect effect of epistemicness on short-term, 95% CI [−0.124, −0.040], but not on long-term events, [−0.027, 0.040], and no significant indirect effect of aleatoriness on either type of event, long-term: [−0.004, 0.008], short-term: [−0.019, 0.006]. In line with what we found in Study 2, this again suggests that the mediating effect of EARS seems to be driven primarily by participants believing that events far into the future are less knowable, rather than more likely to be affected by chance factors.

Study 4—Future Events

In Studies 2 and 3, we tested judgments of forecasters whose accuracy is known, as the outcome of the event has already been realized. However, at the time a prediction is made neither forecasters nor observers know whether a future prediction will turn out to be correct. If observers recognize how epistemic a far future event is, even if the event has not yet occurred, they may discount forecasters' predictions even without knowing what the outcome will be. In Study 4, we test if participants' intuitions about the extent to which an event is influenced by chance factors also influences judgments of forecasters at the time of the prediction itself (i.e., for outcomes that have not yet occurred).

Sample

We recruited 1,283 participants on MTurk, of which 1,054 (53.9% female, 45.0% male, 0.7% another option not listed; $M_{age} = 41.2$ years) passed an attention check that was embedded at the beginning of the study. Participants were paid $0.40 to complete the survey.

Design

Participants were randomly assigned to one of six conditions in a 3 (domain: politics vs. sports vs. entertainment) × 2 (event timing: this year vs. next year) between-subjects design. Participants were told to imagine reading an article published on the day they completed the survey (January 4, 2021) where an expert made a prediction about a future event. The event was either scheduled to take place in the current year (within the next month) or the following year and was in one of three domains (politics, sports, and entertainment). The events were:

> *Politics*—Raphael Warnock will win the Georgia Senate election on [January 5, 2021/ November 8, 2022].
>
> *Sports*—The Kansas City Chiefs will win Super Bowl [55/56] on [February 7, 2021/February 6, 2022].
>
> *Music*—Dua Lipa will win the Grammy for Record of the Year in [2021/2022].

[8] We should note that, although there may be ambiguity about when participants would consider the event itself as having occurred, there should not be any ambiguity about when the *forecast* was made, as we included the exact date of the prediction in our stimuli. The forecasts in the "1 year" condition are all objectively one year (or more precisely, 51 weeks) before the forecasts in the "1 week" condition. As long as participants noted this, the manipulation was likely effective.

Participants then answered the five dependent variables ($\alpha = .93$) and four EARS items ($\alpha = .71$) used in Study 2, presented in counterbalanced order. They also answered a manipulation check question that asked if the events described in the scenario took place in 2021 or 2022. We preregistered that we would only include participants who answered this correctly (87.2% of participants), although our results hold when all participants are included (see Supplement 2 in the online supplemental materials).

Results

Analysis Plan

As preregistered, we excluded those participants who failed to answer the manipulation check correctly and collapsed across prediction domains for our analysis. We used OLS regression to regress the respective measure (forecaster evaluation or EARS) on event timing, including fixed effects for domain.

Main Analysis

Even though the events had not yet occurred and there was no information about the accuracy of the predictions, participants still rated those who made predictions about events a year in the future as less competent ($M = 3.23$, $SD = 1.31$) than those who made predictions about events taking place within the next month ($M = 3.81$, $SD = 1.25$), $b = 0.579$, $t(885) = 6.71$, $p < .001$.

Mediation Analysis

As in prior studies, participants also felt that events occurring a year later were less knowable ($M = 3.10$, $SD = 1.21$) than events occurring within the next month ($M = 3.56$, $SD = 1.24$), $b = 0.436$, $t(885) = 5.46$, $p < .001$. This difference partially mediates our main effect, 95% CI [0.131, 0.296]. These results hold directionally and significantly across all three prediction domains and when including those participants who did not pass the manipulation check (see Supplement 2 in the online supplemental materials). Using both components of the EARS as separate mediators in an exploratory analysis, we find a significant indirect effect of epistemicness, 95% CI [0.161, 0.376] and no significant indirect effect of aleatoriness, [−0.036, 0.008].

Study 5—Prompted Predictions

In the studies presented so far, all of the forecasters in our stimuli provided an unprompted prediction. That is, they made their prediction without being asked to do so. However, oftentimes forecasters may provide a prediction in reaction to being asked by someone or when a newspaper outlet or website wants to release forecasts. Hence, in this study, we test whether our effect holds when a forecaster is being asked to make a prediction about an issue of public interest (e.g., upcoming Senate elections).

Sample

We recruited 1,598 participants on Prolific (51.3% male, 47.2% female, 1.1% another option not listed; $M_{age} = 29.9$ years). Participants were paid $0.60 to complete the survey.

Design

The general setup of this study was similar to that of Study 2. As in Study 2, we used different prediction stimuli and different prediction time frames. Participants were randomly assigned to one of sixteen conditions in a 4 (domain: politics vs. sports vs. investments vs. economics) × 4 (prediction timing: 1 month vs. 1 year vs. 5 years vs. 10 years) between-subjects design.

Participants read a scenario where they were told to imagine reading an article in which an expert made a prediction. In this study, however, they read that the expert was being interviewed and was specifically asked to make a prediction about an event (the 2022 Florida U.S. Senate election, the winner of Super Bowl LVII, the price of bitcoin at the end of February 2023, and the January 2023 U.S. unemployment rate), rather than volunteering an unprompted opinion. Participants were told that the article was written either 1 month, 1 year, 5 years, or 10 years before the event actually happened.

The expert's predictions were:

> *Politics*—Marco Rubio will be elected the US Senator from Florida in the November 2022 elections.
>
> *Sports*—The Kansas City Chiefs will win Super Bowl LVII at the end of the 2022 NFL season.
>
> *Investments*—The price of Bitcoin will be above $20,000 at the end of February 2023.
>
> *Economics*—The US unemployment rate will be below 5% in January 2023.

Participants were told that the prediction was correct and were asked to evaluate the expert using the five dependent variables ($\alpha = .92$) and four EARS measures ($\alpha = .72$) used in Studies 2–4.

Results

Analysis Plan

As preregistered, we collapsed results across the four prediction domains. We used OLS regression to regress the respective measure (forecaster evaluation or EARS) on the number of months in advance the prediction was made (1 vs. 12 vs. 60 vs. 120). The analysis included fixed effects for prediction domain.

Main Analysis

Consistent with the results from our prior studies, there was a negative relationship between prediction timing and forecaster evaluation, $b = -0.001$, $t(1,593) = -2.12$, $p = .035$. In this study, participants' average evaluations remained relatively consistent from 1 month to 5 years in advance (1 month: $M = 5.04$, $SD = 1.07$, 1 year: $M = 5.08$, $SD = 1.10$, 5 years: $M = 5.02$, $SD = 1.09$), and forecasters who made their prediction 10 years in advance were rated the least favorably ($M = 4.91$, $SD = 1.19$). We present the results by stimulus in Supplement 2 in the online supplemental materials.

Mediation Analysis

Participants also rated events to be less epistemic the farther in advance the expert made the prediction. The average EARS score of the event decreased from 4.31 out of 7 ($SD = 1.09$) when the

prediction was made 1 month in advance to 3.50 ($SD = 1.16$) when the prediction was made 10 years in advance, $b = -0.006$, $t(1,593) = -11.04$, $p < .001$. As in prior studies, there is a significant indirect effect of EARS score on evaluations of forecasters, 95% CI [−0.002, −0.001]. We again consider epistemicness and aleatoriness as separate mediators in an exploratory analysis and find that the indirect effect of epistemicness is significant, 95% CI [−0.0027, −0.0016], while the indirect effect of aleatoriness is not, [−0.0002, 0.0003]. This suggests that the perceived knowability is the primary driver of the mediation effect of the EARS.

Taken together, Study 5 shows that our results hold in a scenario where a prediction is explicitly elicited from a forecaster rather than volunteered.

General Discussion

We find that accurate forecasters are not perceived as more competent when they make earlier predictions, and may even be viewed as less competent when they make predictions too early. This is because observers recognize that events far in the future are less knowable.

Consistent with other actor–observer asymmetries (e.g., Critcher et al., 2020; Donnelly et al., 2022; E. E. Jones & Nisbett, 1987), our findings reveal a discrepancy between forecasters' intuitions of how others perceive them and observers' actual perceptions of forecasters. While forecasters may believe that their predictions are based on useful and diagnostic knowledge, observers have less insight into a forecaster's knowledge, process, and motivation, and may rely on situational cues to inform their judgments. Our findings suggest that, if possible, forecasters would benefit from delaying their predictions as long as possible. This would allow them to gather more information about the likely outcome without incurring any negative effects on their perceived competence from waiting.

Is There an "Optimal" Prediction Timing?

Across our studies, we find a consistent negative effect of prediction timing on forecaster evaluations. That is, on average, observers rate forecasters as less competent the earlier they make their predictions. However, it is worth noting that this effect is not necessarily linear. For example, our studies revealed that extremely early forecasts (e.g., years in advance) consistently lower perceptions of competence, but slightly earlier forecasts (e.g., weeks or months in advance) often do not and sometimes directionally increase perceptions of competence. In Study 3, we also found that the perceived knowability of an event moderates the effect of prediction timing on perceived competence. Given this, there may be an "optimal" time to make a prediction where an event is sufficiently knowable but the forecaster still gets extra credit for making an earlier (and therefore, more difficult) prediction. With that said, in our studies, we found no instances where forecasters were seen as significantly more competent for making earlier predictions. Future research may be able to calibrate the precise magnitude of these effects across contexts.

Predictions in Enriched Contexts

In all of the studies presented in the main manuscript, we test a very specific type of prediction. Specifically, we test single, accurate predictions from one forecaster. We do this primarily because this is the simplest and most direct test of our hypotheses. However, we should note that our studies may lack the contextual richness in which predictions are often made. For example, forecasters may make multiple predictions about an event over time or may have a prior history of forecasting different (but similar) events. They may also make predictions alongside other forecasters, offering a useful reference point for how difficult the prediction actually was (Hsee, 1996).

We conducted an initial examination of how contextual aspects may affect judgments of forecasters in Study S3 in Supplement 3 in the online supplemental materials. In this study, we showed participants predictions made by three experts about the 2022 National Football League (NFL) Draft. Each of the three experts provided five predictions, namely which five players they predicted would be the first five selections of the 2022 NFL Draft. The experts made these predictions either 1 week, 6 months, 1 year, or 4 years before the draft occurred. In addition to manipulating the prediction timing, we also manipulated within subjects how many predictions each of the three experts got correct. Participants saw that one expert got all five selections correct, one expert got three out of five correct, and one expert did not get any correct. The results showed that the expert who got all five predictions correct was perceived as equally competent regardless of when they made their prediction. It is possible, however, that this was caused by a ceiling effect due to the comparison to the two less accurate forecasters. In fact, the expert who got all five predictions correct was consistently rated at the very high end of the competence scale, regardless of their prediction timing (between 6.16 and 6.22 out of 7). In contrast, the forecasters who got only three out of five predictions or no prediction correct were perceived as either directionally or significantly[9] less competent the closer to the event they made their prediction. This finding is broadly consistent with the results from a second supplemental study (Study S4 in Supplement 3 in the online supplemental materials) that we conducted where we tested the effect of prediction timing for incorrect predictions. Using similar stimuli as in Studies 2–4, we found that forecasters are generally judged more negatively the *closer* to the event they make an incorrect prediction, although the effect size varies across stimuli (see Supplement 3 in the online supplemental materials).

Taken together, the results from our supplementary studies suggest that both enriching the prediction context and studying the social consequences of incorrect predictions are potentially rich topics for further exploration. In investigating these topics, researchers would need to keep in mind that incorrect predictions can likely not be treated as one construct, as there are different levels of incorrect predictions, including near misses (where the forecaster was almost correct), bad misses (where the forecaster was very far off), and everything in between. Considering the possible moderating impact of a variety of contextual differences, our results should be viewed as an initial exploration of the effect of prediction timing on forecaster evaluations, rather than the final word.

[9] There were typos in two conditions in this study, where the wrong year was listed in the stimuli. However, the description of the prediction timing (e.g., "6 months before the draft") was correct. The results described are significant when including all data and directionally consistent, but not significant, when excluding the two conditions with the errors. See Supplement 3 in the online supplemental materials for all results and analyses.

Other Causes and Effects of Changing Prediction Timing

In our studies, we do not explicitly examine why forecasters make a prediction. We simply tell participants that a prediction was made at a specific time. As a result, we cannot make a normative statement about when forecasters should make predictions. While forecasters who wish to increase their perceived competence may benefit from waiting to make their predictions (since it does not hurt them and they will be able to get more information about the event), there are other reasons why a forecaster may wish to make an earlier prediction. For example, a forecaster may announce a prediction early in order to influence the outcome of an event itself. In the context of a political election, for example, predicting a specific candidate to win might raise enthusiasm for or bring attention to this candidate, possibly increasing their chances of winning. In addition, for many people, speculating may simply be fun and often low cost, possibly resulting in a positive utility of making predictions per se. Similarly, people may enjoy knowing others' predictions, possibly providing pundits with a financial incentive to make premature predictions.

In addition, in our studies, we focus on how changing prediction timing changes perceptions of an event's epistemicness (and how this subsequently affects judgments of competence). However, changing prediction timing could also impact other aspects relevant to judgments of forecasters. For example, by making a premature prediction, forecasters could be seen as violating norms around prediction timing, causing their prediction to be seen as unusual or arrogant. This in turn could lower perceptions of the forecaster's competence or cause participants to rate the forecaster more negatively overall (Gal & Rucker, 2011; Kahneman & Frederick, 2002). In Study S5 in Supplement 3 in the online supplemental materials, we test this possibility. We find that participants directionally, but not significantly, perceive predicting events in the far future to be more of a norm violation ($p = .132$), and perceptions of norm violations also did not mediate the effect of prediction timing on perceptions of competence (i.e., the indirect effect of norm violations included zero; 95% CI [−0.024, 0.200]). Thus, norm violations are unlikely to be the main driver of our effect, although they may still play a minor role. Additionally, forecasts made at different times prior to an event could also imply that different amounts of effort were put into making the prediction. Perhaps such differences could improve (or lower) evaluations of forecasters along dimensions other than competence specifically.

Other Influences of Perceived Epistemicness

We began work on this project with a focus on studying how changing prediction timing changes judgments of forecasters, with perceived epistemicness as a potential mediator. However, as our studies developed, it became clear that perceived epistemicness had a large and consistent effect on judgments of competence, even beyond that of prediction timing. This makes interpreting our mediation results somewhat difficult, considering that the effect of prediction timing is quite small but the effect of epistemicness is quite large. It is likely that there are other attributes of a prediction that may change its perceived epistemicness. For example, the specificity of a prediction (e.g., whether a forecaster simply predicts the winner of a baseball game or its exact score; Kelly & Simmons, 2016) or the number of potential outcomes (Windschitl & Wells, 1998) might make the outcome of an event seem more or less knowable. Forecasters may also benefit from communicating uncertainty in their predictions (Gaertig & Simmons, 2018, 2023), which could better match the content of a prediction to participants' beliefs about how knowable the event is. We believe that there is a great opportunity for more study into these potential attributes and how they influence both perceived epistemicness and judgments of forecasters.

Constraints on Generality

Our studies were conducted using online participants (i.e., from MTurk and Prolific) based in the United States. We chose these samples because they offered us the best way to achieve the sample size necessary to study our effects. Additionally, online samples give us access to a more diverse sample than a U.S. university's behavioral lab or other in-person data collection would. Although online samples are more representative than U.S.-based college samples, they of course may not capture the psychology of those from a diverse range of countries or cultures. We hope that researchers with access to different populations or novel datasets will test our hypotheses in different settings.

Conclusion

In our research, we find that in contrast to people's lay intuitions, forecasters are not perceived as more competent—and are sometimes seen as less competent—the further in advance they make a correct prediction. Observers recognize that events far in the future are less knowable. Forecasters who care about how others evaluate them should take this into account and avoid making premature predictions.

Context

Inspired by how often forecasters claim credit for accurate predictions that seem like simple luck, we were interested to study two consequences of these claims. First, we wanted to know whether forecasters are correct in believing that highlighting such correct predictions is beneficial to their reputation. Second, our prior research investigating the role of epistemicness in predictions (i.e., the extent to which an outcome is knowable in advance) suggested that people do perceive differences in epistemicness across different situations. As such, we set out to examine if observers correctly intuited the role of knowledge and chance in making correct predictions at different points in time and considered this when evaluating the forecasters' skill or expertise. Indeed, we find that the timing of a (correct) forecast affects both the perceived epistemicness of an outcome and the perceived competence of the forecaster. Contrary to their intuitions, forecasters are perceived to be less competent by others when they make predictions far in advance of an event. This is because observers acknowledge that the outcome of an event far in the future is more influenced by chance factors than one that is more immediate. Our findings show that forecasters should consider the reputational consequences of making a premature prediction.

References

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. https://doi.org/10.1037/a0029395

Bagchi, R., & Ince, E. C. (2016). Is a 70% forecast more accurate than a 30% forecast? How level of a forecast affects inferences about forecasts and forecasters. *Journal of Marketing Research*, *53*(1), 31–45. https://doi.org/10.1509/jmr.12.0526

Bambrough, B. (2020, June 8). *Bitcoin could hit $300,000 in five years even without 'institutional adoption.'* Forbes. https://www.forbes.com/sites/

billybambrough/2020/06/08/bitcoin-could-hit-300000-in-five-years-even-without-institutional-adoption/

Bao, H.-W.-S. (2023). *bruceR: Broadly useful convenient and efficient R functions (0.8.10)*. https://cran.r-project.org/web/packages/bruceR/index.html

Beyth-Marom, R. (1982). How probable is probable? A numerical translation of verbal probability expressions. *Journal of Forecasting*, *1*(3), 257–269. https://doi.org/10.1002/(ISSN)1099-131X

Borzello, J. (2021, June 29). *Texas, Oregon moving up in way-too-early top 25 college basketball rankings for 2021–2022*. https://www.espn.com/mens-college-basketball/story/_/page/earlytop25062921/texas-oregon-moving-way-too-early-top-25-college-basketball-rankings-2021-22

Critcher, C. R., Helzer, E. G., & Tannenbaum, D. (2020, March). Moral character evaluation: Testing another's moral-cognitive machinery. *Journal of Experimental Social Psychology*, *87*, Article 103906. https://doi.org/10.1016/j.jesp.2019.103906

Davis, C. (2021, May 3). *2022 Oscars predictions: The collective*. *Variety*. https://variety.com/lists/2022-oscars-predictions-the-collective-top-contenders/

Dhami, M. K., & Mandel, D. R. (2022). Communicating uncertainty using words and numbers. *Trends in Cognitive Sciences*, *26*(6), 514–526. https://doi.org/10.1016/j.tics.2022.03.002

Donnelly, K., Moon, A., & Critcher, C. R. (2022, February). Do people know how others view them? Two approaches for identifying the accuracy of metaperceptions. *Current Opinion in Psychology*, *43*, 119–124. https://doi.org/10.1016/j.copsyc.2021.06.018

Gaertig, C., & Simmons, J. P. (2018). Do people inherently dislike uncertain advice? *Psychological Science*, *29*(4), 504–520. https://doi.org/10.1177/0956797617739369

Gaertig, C., & Simmons, J. P. (2023). Are people more or less likely to follow advice that is accompanied by a confidence interval? *Journal of Experimental Psychology: General*, *152*(7), 2008–2025. https://doi.org/10.1037/xge0001388

Gal, D., & Rucker, D. D. (2011). Answering the unasked question: Response substitution in consumer surveys. *Journal of Marketing Research*, *48*(1), 185–195. https://doi.org/10.1509/jmkr.48.1.185

Gustafson, A., & Rice, R. E. (2019). The effects of uncertainty frames in three science communication topics. *Science Communication*, *41*(6), 679–706. https://doi.org/10.1177/1075547019870811

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, *85*(1), 4–40. https://doi.org/10.1080/03637751.2017.1352100

Howe, L. C., MacInnis, B., Krosnick, J. A., Markowitz, E. M., & Socolow, R. (2019). Acknowledging uncertainty impacts public acceptance of climate scientists' predictions. *Nature Climate Change*, *9*(11), 863–867. https://doi.org/10.1038/s41558-019-0587-5

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, *67*(3), 247–257. https://doi.org/10.1006/obhd.1996.0077

Jones, E. E., & Nisbett, R. E. (1987). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Lawrence Erlbaum Associates.

Jones, Z. (2020, November 3). *AOC for president? Women lead 2024 election odds as bettors favor Alexandria Ocasio-Cortez over Harris, Biden & Trump*. Forbes. https://www.forbes.com/sites/zackjones/2020/11/03/2024-us-presidential-election-odds-women-lead-the-way-as-bettors-favor-alexandria-ocasio-cortez-over-harris-biden–trump/

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin and D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.004

Kelly, T. F., & Simmons, J. P. (2016). When does making detailed predictions make predictions worse? *Journal of Experimental Psychology: General*, *145*(10), 1298–1311. https://doi.org/10.1037/xge0000204

Lefton, H. (2018, August 1). *RBC capital's 10 best tech stocks to buy for 2025*. Kiplinger. https://www.kiplinger.com/slideshow/investing/t058-s001-rbc-s-10-best-tech-stocks-to-buy-for-2025/index.html

Lichtenstein, S., & Newman, J. R. (1967). Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science*, *9*(10), 563–564. https://doi.org/10.3758/BF03327890

McIntyre, H. (2021, May 21). *2022 Grammy preview: This year may be Taylor Swift's biggest yet for nominations*. Forbes. https://www.forbes.com/sites/hughmcintyre/2021/05/21/2022-grammy-preview-this-year-may-be-taylor-swifts-biggest-yet-for-nominations/

Mislavsky, R., & Gaertig, C. (2022). Combining probability forecasts: 60% and 60% Is 60%, but likely and likely is very likely. *Management Science*, *68*(1), 541–563. https://doi.org/10.1287/mnsc.2020.3902

Orr, G., Holmes, K., & Stracqualursi, V. (2022, November 16). *Former President Donald Trump announces a White House bid for 2024*. CNN. https://www.cnn.com/2022/11/15/politics/trump-2024-presidential-bid/index.html

Peek, D. (2020, November 5). *Biden win would leave GOP poised for 2024 comeback* [Text]. The Hill. https://thehill.com/opinion/white-house/524615-biden-win-would-leave-gop-poised-for-2024-comeback

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. https://doi.org/10.1002/bdm.460

Radzevick, J. R., & Moore, D. A. (2011). Competing to be certain (but wrong): Market dynamics and excessive confidence in judgment. *Management Science*, *57*(1), 93–106. https://doi.org/10.1287/mnsc.1100.1255

Tannenbaum, D., Fox, C. R., & Ülkümen, G. (2017). Judgment extremity and accuracy under epistemic vs. aleatory uncertainty. *Management Science*, *63*(2), 497–518. https://doi.org/10.1287/mnsc.2015.2344

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. https://doi.org/10.1016/j.jesp.2008.04.006

Van Zant, A. B. (2022). Strategically overconfident (to a fault): How self-promotion motivates advisor confidence. *Journal of Applied Psychology*, *107*(1), 109–129. https://doi.org/10.1037/apl0000879

Watts, D. J., & Rothschild, D. M. (2017, December 5). *Don't blame the election on fake news. Blame it on the media*. Columbia Journalism Review. https://www.cjr.org/analysis/fake-news-media-election-trump.php

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin*, *25*(9), 1115–1125. https://doi.org/10.1177/01461672992512005

Windschitl, P. D., & Weber, E. U. (1999). The interpretation of "likely" depends on the context, but "70%" is 70%—right? The influence of associative processes on perceived certainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1514–1533. https://doi.org/10.1037/0278-7393.25.6.1514

Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology: Applied*, *2*(4), 343–364. https://doi.org/10.1037/1076-898X.2.4.343

Windschitl, P. D., & Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology*, *75*(6), 1411–1423. https://doi.org/10.1037/0022-3514.75.6.1411

Yaniv, I., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology: General*, *124*(4), 424–432. https://doi.org/10.1037/0096-3445.124.4.424

Received August 3, 2022
Revision received July 17, 2023
Accepted August 5, 2023 ■