Journal of Experimental Social Psychology 113 (2024) 104619

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

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

Uncertainty, expertise, and persuasion: A replication and extension of Karmarkar and Tormala (2010)☆

Erik Løhre [a,*], Subramanya Prasad Chandrashekar [b], Lewend Mayiwar [a], Thorvald Hærem [a]

[a] Department of Leadership and Organizational Behaviour, BI Norwegian Business School, Oslo, Norway
[b] Department of Psychology, NTNU – Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Keywords:
Attitudes
Uncertainty
Confidence
Expertise
Replication
Registered report

ABSTRACT

If you are trying to persuade someone, expressing your opinion with certainty intuitively seems like a good strategy to maximize your influence. However, Karmarkar and Tormala (2010) found that the effectiveness of this tactic depends on expertise. In three experiments, Karmarkar and Tormala found support for an incongruity hypothesis, whereby non-expert sources can gain interest and influence by expressing certainty, while expert sources can increase persuasion by expressing uncertainty. In this Registered Report, we conducted a high-powered (N = 1018) direct replication of Experiment 2 by Karmarkar and Tormala (2010). In a consumer behaviour context, the original study examined whether source expertise moderated the positive effect of source certainty on the persuasive impact of a restaurant recommendation. The present replication failed to find support for the incongruity hypothesis, $\eta_p^2 = 0.00$ [0.00, 0.02]: expressing certainty had a positive but non-significant effect for non-experts, $d = 0.10$ [−0.10, 0.34], and a positive effect for experts, $d = 0.28$ [0.03, 0.52]. Instead, the results supported the competing *confidence heuristic* hypothesis that expressed certainty would have a positive effect on persuasion, irrespective of source expertise, $d = 0.18$ [0.01, 0.36]. Extending the original work, we (1) controlled for the reason given for (un)certainty, and (2) examined need for closure as a potential individual difference moderator. The results indicated robust support for the confidence heuristic $d = 0.25$, [0.12, 0.37], but neither reason for (un)certainty nor need for closure moderated the effect as hypothesized. All materials, data, and code are available on: https://osf.io/hbjyv/.

In a complex world, we often look to experts to guide our decisions and opinions (Ahluwalia, Edelen, Qureshi, & Etchegaray, 2021; Harvey & Fischer, 1997). This is the case for topics ranging from consumer choices (Which restaurants do the critics think are good? Which phone gives me the best value for money?) to scientific and political issues (Should I socially distance from others during the virus outbreak? Which policy gives the best chance of mitigating climate change?). It is crucial to understand which factors can make experts, real as well as fake, more persuasive.

Uncertainty is a central topic in this regard. Experts have higher knowledge about their domain than lay people and can therefore often be more certain about their assertions. But when issues are complex or new, even experts cannot claim certainty, potentially leading to a disconnect between how certain experts can be and how certain lay people expect them to be (Corner, Lewandowsky, Phillips, & Roberts, 2015). However, a 2010 article by Karmarkar and Tormala made the intriguing discovery that experts can be more persuasive if they express uncertainty. This is an important finding that allows experts to be transparent about uncertainty, but still retain or even increase their persuasiveness.

In the current registered report, we conducted an independent close replication of Karmarkar and Tormala's (2010) study that found that uncertainty can increase persuasiveness for experts. Furthermore, we extended the study by controlling for a potential confound in the original study design, namely reasons for certainty/uncertainty, and by examining individual differences in need for closure (Webster & Kruglanski, 1994) as a potential moderator of the effect of certainty vs. uncertainty.

1. Uncertainty and persuasiveness

Before discussing effects of expressing uncertainty, it is necessary to

☆ This paper has been recommended for acceptance by Ed Hirt.

* Corresponding author at: Department of Leadership and Organizational Behaviour, BI Norwegian Business School, 0442 Oslo, Norway.
E-mail address: erik.lohre@bi.no (E. Løhre).

https://doi.org/10.1016/j.jesp.2024.104619
Received 20 February 2024; Received in revised form 9 April 2024; Accepted 12 April 2024
Available online 18 April 2024

define what is meant by uncertainty. This is not straightforward, as the construct of uncertainty has been studied and conceptualized in many different ways (Smithson, 2008). Some theorists distinguish different types of uncertainty based on the reason for uncertainty, such as natural variability, disagreements, or lack of knowledge (Gustafson & Rice, 2020). Others distinguish between different sources of uncertainty, such as whether the uncertainty is internal (a subjective feeling based for instance on lack of knowledge) or external (an objective aspect of an event based on relative frequencies or on propensities; see e.g. Kahneman & Tversky, 1982; Løhre & Teigen, 2016; Teigen, 1988). A similar dichotomy uses the labels epistemic ("knowable") and aleatory ("random") uncertainty (Ülkümen, Fox, & Malle, 2016; Walters, Ülkümen, Tannenbaum, Erner, & Fox, 2023), referring to whether the uncertainty is due to missing information or expertise about an event that is knowable in principle (epistemic uncertainty) vs. due to the stochastic/random nature of a class of events (aleatory uncertainty).[1]

These distinctions are important, as the effects of expressing uncertainty on observer perceptions may depend on the type of uncertainty and/or on how uncertainty is conceptualized (Gaertig & Simmons, 2018; Gustafson & Rice, 2020; Løhre & Teigen, 2023; van der Bles, van der Linden, Freeman, & Spiegelhalter, 2020), although the research on this topic is scarce. In the present investigation, we focus on the concept of attitude certainty (Tormala & Rucker, 2007), which can be seen as a type of internal or epistemic certainty; a subjective feeling regarding the conviction with which one holds an attitude.

If the aim is to persuade or influence others, expressing certainty about an opinion seems like an intuitively appealing strategy. There is a stream of research supporting this view: people are more likely to follow advice from sources with high confidence (Bonaccio & Dalal, 2006; Radzevick & Moore, 2010; Sniezek & Van Swol, 2001), to judge leaders as more competent when they express certainty rather than uncertainty (Løhre & Teigen, 2023), and eyewitnesses who express higher confidence are seen as more credible (Tenney, MacCoun, Spellman, & Hastie, 2007) and have greater influence on courtroom decisions (Brewer & Burke, 2002). Price and Stone (2004) have called this a confidence heuristic: people assume that those who are highly confident are confident because they know what they are talking about, and hence confidently expressed opinions are more persuasive than humbly or uncertainly expressed opinions.

According to the confidence heuristic, experts should express (internal) certainty if they wish to influence others. In contrast to this view, Karmarkar and Tormala (2010) proposed that for experts, it can pay off to express uncertainty. The reasoning is based on an informational incongruity perspective: when two different salient persuasion variables are incongruent with each other, it can lead to increased engagement and involvement with a message (Maheswaran & Chaiken, 1991; Petty & Cacioppo, 1986). Thus, assuming that the message contains good arguments, when the source is a non-expert, people are more persuaded by certainty than by uncertainty. However, for expert sources, people are surprised to hear them expressing uncertainty, and are thus more engaged with the message and more persuaded.

In a recent theoretical article, Karmarkar and Tormala's (2010) findings have been cited as an example of an "act of receptiveness" (Hussein & Tormala, 2021), that is, a behaviour that signals openness to new information and opposing viewpoints. According to Hussein and Tormala, such behaviours can boost persuasion by increasing involvement and enhancing source perceptions, especially when they stem from expert or high-status sources.

The main hypothesis investigated in Karmarkar and Tormala (2010) can be summarized as follows:

H1a. For non-expert sources, source certainty will have a positive effect on persuasion, while for expert sources, source certainty will have a negative effect on persuasion.

In other words, this *incongruity hypothesis* states that people will be more persuaded by non-experts expressing their opinions with certainty rather than uncertainty, but will be more persuaded by experts expressing their opinions with uncertainty rather than certainty. This can be described as an interaction effect of expertise and certainty on persuasion. The findings reported by Karmarkar and Tormala are consistent with this hypothesis. In Experiment 1, source expertise and expressed certainty was varied between subjects, and participants rated a message as more surprising when expertise and certainty were incongruent (e.g., high expertise, low certainty). In Experiment 2, participants reported greater involvement with the message and had more favorable attitudes when expertise and certainty were incongruent. In other words, they were more persuaded by uncertain experts and certain novices. Finally, in Experiment 3, argument quality was manipulated in addition to certainty and expertise, and there were stronger effects of argument quality on attitudes when expertise and certainty were incongruent. Overall, Karmarkar and Tormala's findings indicate that experts may benefit from expressing their uncertainty. Table 1 provides an overview of the main findings for the three experiments reported in the original article.

2. Choice of replication target: Karmarkar and Tormala (2010), Experiment 2

There are several reasons we believe that a close, independent replication of Karmarkar and Tormala (2010) would have high value (Isager et al., 2023).

First, the findings are important. It is crucial for experts to know whether and how they should communicate uncertainty while retaining their credibility. This is evidenced in several recent articles concerning uncertainty communication related to different scientific topics like climate change and Covid-19 (Kause et al., 2021; Kelp, Witt, & Sivakumar, 2022; Løhre, 2018; van der Bles et al., 2020).

Second, the findings have had large impact. As of April 2024 Karmarkar and Tormala's article has been cited 346 times (Google Scholar), including in popular science books (Grant, 2021), extending the reach outside of academic circles.

Third, the findings are relatively counterintuitive, and contrast with previous research on the confidence heuristic (Price & Stone, 2004).

Fourth, the original study had some limitations. As shown in Table 1, the sample sizes were relatively small, giving the study low statistical power to detect anything other than large effects. This is also evidenced in the wide confidence intervals for the effect sizes, which range from very small to very large.

Fifth, to the best of our knowledge, there has not been any direct replication of Karmarkar and Tormala's study. In accordance with the recent focus on the importance of replication studies to assess reproducibility and generalizability of findings (Asendorpf et al., 2013; Open Science Collaboration, 2015; Zwaan, Etz, Lucas, & Donnellan, 2018) we decided to revisit these findings. We chose Experiment 2 as our replication target, as this provides the most straightforward demonstration of the core phenomenon of interest, namely that experts can be more persuasive if they express uncertainty.

3. Competing hypothesis: confidence heuristic

As explained above, many studies have documented a positive effect of expressing certainty rather than uncertainty on for instance advice taking and perceived competence (e.g., Bonaccio & Dalal, 2006; Løhre & Teigen, 2023; Price & Stone, 2004). Similarly, some studies of expert communication of risks argue that expressing uncertainty may negatively influence public attitudes toward scientists and their claims

[1] For more details about the similarities and differences between the internal vs. external and epistemic vs. aleatory conceptualizations, see Teigen and Løhre (2017), Fox and Ülkümen (2017), and Juanchich, Gourdon-Kanhukamwe, and Sirota (2017).

Table 1
Key findings for the three experiments reported in Karmarkar and Tormala (2010).

Experiment	N	IVs	Main DV of interest[a]	Expertise x certainty interaction[b]	Simple effect of certainty, experts[c]	Simple effect of certainty, non-experts	Main finding
1	105	Expertise, certainty	Expectancy violation	$F(1,101) = 7.55$, $p < .008$, $\eta_p^2 = 0.07$ [0.01, 0.18]	$F(1,101) = 3.17$, $p < .08$, $d = -0.35$ [−0.75, 0.04]	$F(1,101) = 4.43$, $p < .04$, $d = 0.42$ [0.02, 0.81]	More surprising with non-expert certainty and expert uncertainty
2	68	Expertise, certainty	Attitudes	$F(1,64) = 9.70$, $p < .004$, $\eta_p^2 = 0.13$ [0.02, 0.29]	$F(1,64) = 4.94$, $p < .03$, $d = -0.56$ [−1.05, −0.05]	$F(1,64) = 4.77$, $p < .04$, $d = 0.55$ [0.04, 1.04]	More persuasive with non-expert certainty and expert uncertainty
3	140	Expertise, certainty, argument quality	Attitudes, thought favorability	NA[d]	NA	NA	Qualitatively the same pattern as before for strong arguments: more persuasive with non-expert certainty and expert uncertainty

[a] We focus on the dependent variables that are most important for the main hypothesis. Other dependent variables were also reported by Karmarkar and Tormala (2010).

[b] Effect sizes with 95% confidence intervals for both the interaction and the simple effects are computed based on *F*-values and degrees of freedom using the *easystats* package in R (Lüdecke, Makowski, Ben-Shachar, Patil, & Wiernik, 2022); code and output is provided on https://osf.io/hbjyv/?view_only=4fa551dff0574e0b9c24f80c1d18cf4a.

[c] For the simple effects, Cohen's *d* is coded such that a positive sign indicates higher score on the attitude measure in the high certainty than in the low certainty condition, while a negative sign indicates a lower score on the attitude measure in the high vs. the low certainty condition.

[d] For Experiment 3, a three-way interaction is reported, followed by two separate two-way interactions. The two-way interactions are split by expertise (i.e., the interaction between argument quality and certainty are reported separately for experts and non-experts). This (plus the lack of access to the original data) means that we cannot easily compute the relevant statistics.

(Gustafson & Rice, 2020). Thus, a competing hypothesis to Karmarkar and Tormala's interaction hypothesis is that expressed certainty would have a general positive effect on persuasion, independent of expertise. We formulate this competing hypothesis of a main effect of certainty on persuasion as follows:

H1b. People will report more positive attitudes to a certain recommendation over an uncertain recommendation.

4. Extensions

4.1. Main extensions

We extended the original article in two ways: (1) controlling for a potential confound (reasons for [un]certainty), and (2) investigating need for closure as a potential individual difference moderator.

Firstly, a close reading of the original scenario revealed a potentially important difference between the certainty and uncertainty conditions. Below are the relevant passages from the original scenario, with the uncertainty condition in brackets:

> "I am certain [Of course, I can't be certain] that the chef has done [can do] all of the dishes on the menu as well as these. (…) Having eaten there for dinner, I can confidently [Having eaten there only once, I don't have complete confidence in my opinion, but I suppose I would] give Bianco a rating of 4 (out of 5) stars."

Note that in the uncertainty condition, a good reason is given for the reviewer's uncertainty, namely the lack of data/low number of observations. It is reasonable to express low certainty in your opinion of the quality of a restaurant that you only went to once and where you only tasted a couple of dishes. This lack of data could be associated with both epistemic/internal uncertainty (due to lack of knowledge) as well as aleatory/external uncertainty (due to random variations in performance), although the reviewer communicates this as internal uncertainty (*I* can't be certain, *I* don't have complete confidence). In the certainty condition, however, no reason is given for the reviewer's certainty. This is a potential confound, as uncertainty with a (good) reason is compared to certainty without a reason.

Thus, while we retained and compared the original conditions to ensure a direct replication, we also as an extension added reason for (un)certainty as a factor, so that uncertainty with a reason could be compared to certainty with a reason, and vice versa. Importantly, by accounting for reason were able to empirically test the robustness of the predictions of the incongruity hypothesis. We kept the original reason for (internal) uncertainty (lack of data) and compared this with what we believe is a plausible reason for the reviewer expressing (internal) certainty, namely that the performance in the restaurant was on a high level despite external challenges (difficult dishes and a busy night).

We expected that reason might be of importance. Mainly, we predicted more positive attitudes when a reason is given, for both certainty and uncertainty, leading to the following extension hypothesis:

H2. People will report more positive attitudes toward a recommendation with a good reason for (un)certainty than toward the recommendation with an absence of a good reason.

Secondly, we investigated a potential moderator, namely individual differences in need for closure (Webster & Kruglanski, 1994). In their recent article, Hussein and Tormala (2021) argue that effects of "acts of receptiveness" (of which expressing uncertainty is an example) may differ depending on individuals' dispositions, for instance their tolerance of uncertainty or ambiguity. It is well known that people in general are ambiguity averse (Ellsberg, 1961; see Bühren, Meier, & Pleßner, 2023 for a recent review). The need for closure scale (NFCS; Roets & Van Hiel, 2011) is a measure that captures individual differences in preferences for order and structure, affective discomfort with ambiguity, and the desire for secure or stable knowledge. Previous research has found that those who have a higher need for closure are more susceptible to influence tactics that reduce ambiguity (Kardes, Fennis, Hirt, Tormala, & Bullington, 2007), and are more motivated to resolve open and ambiguous decision situations by relying on decision strategies that bypass the decision making process (e.g., delegating the decision or choosing a default option; Otto, Clarkson, & Kardes, 2016). Thus, our final extension hypothesis is the following:

H3. The influence of the certainty of a recommendation on people's attitudes toward the recommendation is moderated by need for closure (NFC) such that positive attitudes to a certain recommendation over an uncertain recommendation will be stronger among people with higher NFC than lower NFC.

Table 2 provides an overview of the replication and extension hypotheses.

Table 2
Hypotheses investigated in the current study.

Hypothesis #	Description	Replication/ extension
1a	Source certainty will have a positive effect on persuasion for non-expert sources, while for expert sources, source certainty will have a negative effect on persuasion.	Replication
1b	People will report more positive attitudes to a certain recommendation over an uncertain recommendation	Competing hypothesis
2	People will report more positive attitudes to a recommendation with a reason for (un)certainty than to a recommendation with no reason for (un)certainty	Extension
3	The influence of the certainty of a recommendation on people's attitudes toward the recommendation is moderated by NFC such that positive attitudes to a certain recommendation over an uncertain recommendation will be stronger among people with higher NFC rather than lower NFC	Extension

4.2. Additional exploratory questions

The incongruity hypothesis proposed by Karmarkar and Tormala (2010) suggests that the interaction effect between source certainty and expertise occurs because of the incongruity between these two salient persuasion variables. This proposed mechanism was investigated in their Experiment 1, which showed that participants rated it as more surprising and unexpected when an expert expressed uncertainty and a non-expert expressed certainty, rather than the other way around. As an exploratory test of the same mechanism, we included the two questions measuring expectancy violation that were used by Karmarkar and Tormala in Experiment 1.

A final set of exploratory questions were included to probe whether the participants perceived the reviewer's (un)certainty as internal vs. external (Løhre & Teigen, 2016, 2023) and/or as epistemic/aleatory (Ülkümen et al., 2016). These questions were included since previous studies indicate that responses to (un)certainty may differ depending on the type of uncertainty. For instance, in Løhre and Teigen (2023) there was a larger difference between responses to internal certainty vs. uncertainty as compared to external certainty vs. uncertainty. Also, including these questions allow us to probe whether the different reasons provided for certainty and uncertainty differed with respect to the perceived type of uncertainty.

5. Open science statement

The study was pre-registered on the Open Science Framework (OSF) after in-principle acceptance of the registered report, and data collection commenced shortly after. All study materials, code, and collected data are provided on: https://osf.io/hbjyv/.

The study received approval from Sikt - Norwegian agency for shared services in education and research (#391346) and from the ethical review board at BI Norwegian Business School (#023). Furthermore, we note that all measures, manipulations, and exclusions for this investigation are reported (either in the main text or in the supplement), and that the data collection wascompleted before the analyses.

6. Method

6.1. Power analysis

The power analyses were based on three criteria. The first and most important consideration was to ensure sufficient power to detect the effect sizes reported in Experiment 2 of Karmarkar and Tormala (2010) in the conditions directly replicating the original experiment. The results of a power analysis computed with the *Superpower* package in R (Lakens & Caldwell, 2021) indicated that a sample size of 304 participants would have a power >0.95 to detect the effect size of the interaction between source expertise and source certainty on the attitude measure (Cohen's $f = 0.39$ / $\eta_p^2 = 0.13$) as reported in Experiment 2 of Karmarkar and Tormala (2010). Thus, a sample size of 304 should be sufficient to conduct a high-powered direct replication of the original experiment.

The second criterion concerned the extension involving the reason for (un)certainty. Adding the new conditions (certainty with reason and uncertainty without reason) changes the design from the original 2 × 2 to a 2 × 2 × 2 design. Hence, we conducted a power analysis (using G*Power, Faul, Erdfelder, Lang, & Buchner, 2007) with an aim to detect a small to medium interaction effect size of Cohen's $f = 0.17$ in a 2 × 2 × 2 between-subjects experimental design at 0.95 power ($\alpha = 0.05$). The analysis suggested a total sample size of 452 or more.

The third criterion was to estimate a total sample size to account for the second extension involving the potential moderating effect of need for closure (measured at the individual level) on the relationship between certainty and attitude. We conducted a power analysis in G*Power with an aim to detect a small interaction effect size of $f^2 = 0.02$ in a multiple regression analysis accounting for the interaction effect of Need for closure × Certainty (High vs. Low) on the attitudes measure (dependent variable) at 0.95 power ($\alpha = 0.05$). The analysis suggested a total sample size of 652 or more.

Based on these three power analyses, our objective was to achieve a total sample size of 1000 participants. This planned sample size fulfilled all three of the criteria explained above. Most importantly, the total sample size ensured that around 500 participants would be allocated to the subset of conditions that directly replicate the original Karmarkar and Tormala Experiment 2 design, exceeding the recommended sample size of 304. For a detailed overview of power analysis simulations, please see the supplementary file.

6.2. Participants

A total sample of 1018 participants from the US via Prolific completed the study. The sample included 487 women, 509 men, 10 non-binary, and 11 unreported ($M_{age} = 40.12$ years, $SD_{age} = 14.03$). We used standard Qualtrics spam prevention measures (e.g., reCAPTCHA, prevent multiple submission, prevent ballotstuffing, bot detection). We paid participants according to Prolific's guidelines for fair pay (currently £9.00/$12.00 per hour). For a survey taking about 9 min (based on 20 pretest-responses which were not included in the analyses), this gives a payment of £1.35. The current sample is compared with the sample from the original study in Table 3.

Table 3
Comparison of the sample in the original study and in the replication.

	Karmarkar and Tormala (2010), Exp. 2	Replication
Sample size	68	1018
Geographic origin	US American	US American
Population	Undergraduates	Prolific participants
Gender	Not reported	487 women, 509 men, 10 non-binary, and 11 unreported
Median age	Not reported	37.0 years
Average age	Not reported	40.12 years
Standard deviation age	Not reported	14.03
Age range	Not reported	18–81 years
Medium (location)	Computer (lab)	Computer (online)
Compensation	Yes, but amount not reported	£1.35
Year	Probably 2007 or 2008	2023

6.3. Design and procedure

Participants responded to the survey online using Qualtrics. After providing their informed consent and reading a short introduction to the study, participants confirmed that they had understood their task. They then proceeded to the scenario, which was a direct replication of the scenario used by Karmarkar and Tormala (2010), a review of an Italian restaurant in the US. Following the original article, participants were randomly assigned to different conditions in a 2 × 2 design with source certainty (low vs. high) and source expertise (low vs. high) as between-subjects factors. As an extension, we included reason for (un)certainty (no reason vs. reason) as an additional between-subjects factor. Note however, that in our analysis we first compare only the conditions that were included in the original study (i.e., certainty with no reason vs. uncertainty with a reason for experts vs. non-experts), and that the new conditions (certainty with a reason vs. uncertainty with no reason) are only included in later analyses. This was done to ensure that our first analysis is a direct replication of the analysis in the original study.

As another extension, we included need for closure (Webster & Kruglanski, 1994) as a potential individual difference moderator.

A detailed description of the procedure and materials are provided in the section about "Materials and scales" in the supplementary materials, which also contains information about some necessary adjustments to the design and procedure of the original study (see Table S5 for an overview). Fig. 1 shows a screenshot of the restaurant review as displayed in the low expertise, low certainty with reason (replication) condition, and Table 4 describes how statements in the texts differed in the experimental conditions.

6.4. Independent variables

Manipulation of source expertise. In different conditions, the restaurant review stemmed from a source with low or high expertise. The reviewer was described as "a networks administrator" in the low expertise condition and as "a nationally renowned food critic" in the high expertise condition, and the opening sentence of the review stated that the reviewer usually ate fast food (low expertise) or had eaten at most of the Italian places nearby (high expertise), see also Table 4.

Manipulation of source certainty. As demonstrated in Table 4, source certainty was manipulated in three different places in the text: in the headline, in the middle of the text, and in the conclusion. In the conditions directly replicating Karmarkar and Tormala (2010), source certainty was confounded with reason, with low certainty including a reason, while high certainty did not include a reason. We thus included two new conditions: low certainty without a reason and high certainty with a reason. To briefly summarize the manipulations of source certainty, the headline either gave a "tentative" or a "confident" 4 out of 5 stars; the middle of the text contained a short passage where the reviewer stated "I am certain" or "I can't be certain"/"I am not certain" that the quality of all dishes in the restaurant were equally high as the ones he had tasted; and in the conclusion the reviewer stated "I don't have complete confidence" or "I can confidently" give a 4 star rating.

Manipulation of reason for (un)certainty. In the original study, low certainty was explained with reference to the small sample (only tasting a few dishes, only visiting the restaurant once). As shown in Table 4, in the new "low certainty without a reason"-condition, these references to the "lack of data" for an attitude was removed. In the new "high certainty with reason"-condition, we explained the certainty of the reviewer's attitude by having him point out that the restaurant staff delivered high quality even if the task was difficult (the dishes were challenging to make, the performance was good on a busy night). Note that this was not a mirror image of the "low certainty with reason"-condition in the original, i.e., we did not explain certainty with a reference to a large amount of data. This would have introduced experience (which is related to expertise) as a new potential confound. The point was not to have symmetrical explanations of low and high certainty, or to be able to distinguish between the effects of different types of reasons, but rather to see whether a plausible (and presumably good) reason for expressing internal (un)certainty about an opinion influenced the results.

Individual difference moderator: Need for closure. The potential moderating individual difference variable need for closure (Webster & Kruglanski, 1994) was measured at the end of the survey, right before attention checks and demographic questions. We opted for the brief 15-item revised need for closure scale (NFCS) described by Roets and Van Hiel (2011). The scale consists of 15 items such as "I don't like situations that are uncertain" and "When I have made a decision, I feel relieved", with participants rating their agreement with these statements on a scale from *Completely disagree* (1) to *Completely agree* (6). See supplementary materials for an overview of all 15 items in the NFCS.

6.5. Dependent variables

Attitudes. The main dependent variable was participants' attitudes after reading the restaurant review. As in the original article, this was measured by asking: "After reading this review, what are your attitudes towards Bianco?", which was rated on three bipolar 9-point scales, *Negative* (1) to *Positive* (9), *Bad* (1) to *Good* (9), and *Unfavorable* (1) to *Favorable* (9). In the analysis, we combined these three ratings into an average attitude score.

Involvement. In Karmarkar and Tormala (2010), involvement with the review functioned as a mediator variable explaining the interaction of source certainty and expertise on attitudes. Although this was not our main interest, we included the same two questions, namely, "How involved did you feel with the review of Bianco?" and "How interested were you in the restaurant review?", with nine-point scales from *Not involved [interested] at all* (1) to *Very involved [interested]* (9). These two items were combined into an average involvement score.

Manipulation checks: perceived certainty, expertise, and reason for (un)certainty. As in the original study, two questions probing participants' impressions of the certainty and expertise of the reviewer were included. First, participants answered "How certain was the author of the review of his assessment of Bianco?", on a scale from *Not certain at all* (1) to *Extremely certain* (9). Next, they answered "What level of expertise did the author of the review have about restaurants?", on a scale from *Not expert at all* (1) to *Very expert* (9). Additionally, as a manipulation check of our new manipulation of reason for (un)certainty, participants were asked "Did the author of the review have a good reason for being [un]certain about his assessment of Bianco?", on a scale from *Did not have a good reason at all* (1) to *Had a very good reason* (9). Note that participants in high certainty conditions were asked about the reason for certainty while participants in low certainty conditions were asked about the reason for uncertainty.

6.6. Exploratory measures

Expectancy violation. After answering the manipulation check questions, participants received two questions taken from Karmarkar and Tormala's (2010) Experiment 1, measuring the degree to which participants perceived the review as surprising and unexpected, on scales from *Not at all surprising/unexpected* (1) to *Extremely surprising/unexpected* (9).

Perceived type of (un)certainty. We included also two questions adapted from Study 5 in Løhre and Teigen (2023), measuring whether the uncertainty was perceived as external or internal and four questions adapted from Ülkümen et al.'s (2016) epistemic-aleatory rating scale (EARS). First, participants rated their agreement on scales from *Completely disagree* (1) to *Completely agree* (7) with statements that the reviewer expressed that he was [un]certain about his assessment of Bianco "…based on objective facts that would be apparent to other people" (external uncertainty) and "…because of a subjective feeling he had" (internal uncertainty). Second, participants indicated to what

Fig. 1. Screenshot of the restaurant review scenario from the low expertise, low certainty with reason (replication) condition.

Table 4
Overview of manipulations of independent variables.

Source expertise (replication)	Low	High
	The review you will see was written by Daniel Christiansen, who is a networks administrator at a community college.	The review you will see was written by Daniel Christiansen, who is a nationally renowned food critic and a regular contributor to the food and dining section of a major area newspaper.
	I usually end up eating out at fast food places, but last night I was invited by a friend to try Bianco, an elegant mid-priced restaurant on the peninsula that just opened a few months ago.	I have had the opportunity to eat at most of the Italian places nearby, but last night I was invited by a friend to try Bianco, an elegant mid-priced restaurant on the peninsula that just opened a few months ago.
Source certainty (replication)	**Low (with reason)**	**High (without reason)**
	Bianco – a tentative 4 out of 5	Bianco – a confident 4 out of 5
	Of course, I can't be certain that the chef can do all of the dishes on the menu as well as these.	I am certain that the chef has done all of the dishes on the menu as well as these.
	Having eaten there only once, I don't have complete confidence in my opinion, but I suppose I would give Bianco a rating of 4 (out of 5) stars.	Having eaten there for dinner, I can confidently give Bianco a rating of 4 (out of 5) stars.
Source certainty controlling for reason (extension)	**Low (without reason)**	**High (with reason)**
	Bianco – a tentative 4 out of 5	Bianco – a confident 4 out of 5
	However, I am not certain that the chef has done all of the dishes on the menu as well as these.	I am certain that the chef has done all of the dishes on the menu as well as these, since these dishes are quite challenging to make.
	I don't have complete confidence in my opinion, but I suppose I would give Bianco a rating of 4 (out of 5) stars.	Seeing that the restaurant staff performed so well on a busy night, I can confidently give Bianco a rating of 4 (out of 5) stars.

extent they got the sense that the reviewer believes the performance of a restaurant is "knowable in advance given enough information" and "is something that becomes more predictable with additional knowledge or skills" (epistemic items), or "is determined by chance factors" and "is something that has an element of randomness" (aleatory items), on a scale from *Not at all* (1) to *Very much* (7).

6.7. Other questions

Attention checks. After completing the scenario and the need for closure scale, participants were given a simple attention check. A brief text explained that it is important for the research that participants give their full attention to the instructions and that to detect those who just skim through it, they should answer "Sports". Then, they were asked "Based on the text you read above, what was the topic of this survey?", with four alternatives, Politics, Climate change, Criminal investigations, and Sports. Another question asked "How serious were you in filling out this questionnaire?" on a scale from *Not at all* (1) to *Very much* (5). In the same block of the survey, participants were also allowed to describe what they thought the purpose of the study was (open ended response of one sentence) and rated their understanding of the English used in the study on a 9-point scale from *Very bad* to *Very good*.

Demographics. At the end of the survey, participants provided their age (in years), sex, education level, and country of residence.

6.8. Replication closeness

Deviations. Some minor adjustments were made to the procedure used in the original study. These are explained in more detail in the supplementary materials (see Table S5). In brief, since we did not have access to the full materials of the original study, we wrote a study introduction based on the description in Karmarkar and Tormala (2010). Furthermore, while the original stated that the review was taken from an "online journal", we described the review as "published on a website that collects reviews of restaurants from food critics as well as ordinary restaurant guests". Additionally, we updated the stated price for the restaurant visit from $25 per person in the original to $60, which is a more realistic price per person in a mid-priced Italian restaurant currently.

Evaluation of replication closeness. We evaluated replication closeness based on the criteria described by LeBel, McCarthy, Earp, Elson, and Vanpaemel (2018). As shown in Table 5, we classify this as a very close replication.

Evaluation of replication findings. We evaluated the replication

Table 5
Evaluation of replication closeness.

Design facet	Replication	Details
Effect/hypothesis	Same	
IV construct	Same	
DV construct	Same	
IV operationalization	Same	
DV operationalization	Same	
Population (e.g., age)	Different	The original study used US undergraduates as participants, while the replication recruits from an online platform with a more diverse sample from the US. The average age will be higher, the range of age will be larger, and the backgrounds will be more diverse.
IV stimuli	Same	
DV stimuli	Same	
Procedural details	Similar	Several minor adjustments were made. Some due to missing information from the original study, others to adapt to current situations (e.g., current prices for a restaurant visit), others again to ensure data quality (e.g., attention checks). Extensions were also added.
Physical settings	Similar	Original was conducted via computer in a lab, replication is conducted via computer online.
Contextual variables	Different	Trust in experts may or may not differ between 2007/8 and 2023; people may have more experience with expert uncertainty, for instance after the COVID pandemic.
Replication classification	Very close replication	

findings based on the criteria reported by LeBel et al. (2018); LeBel, Vanpaemel, Cheung, and Campbell (2019). This entails comparing the confidence interval for the replication effect size with the original effect size estimate, and reporting (1) whether there is a signal (i.e., whether the replication effect size excludes zero), (2) whether the effect is consistent (i.e., whether the confidence interval includes the point estimate of the original effect size or is smaller/larger/in the opposite direction), and (3) whether the replication effect size estimate is more or less precise than the original estimate.

6.9. Data analysis plan, exclusions, and handling of outliers

We pre-registered a detailed data analysis plan, which we followed closely in our reported results. The full data analysis plan as reported in our Stage 1 manuscript can be found in the supplement.

The study included an initial comprehension check, an attention check, and a question about seriousness when responding. Even though these measures were included, we retained all participants, but as a robustness check, we also ran the analyses after excluding those who indicated they did not understand the instruction, failed the attention check, indicated that they were not serious when responding (i.e., rated their seriousness as below 3 on a 5-point scale), those who correctly guessed the main purpose of the study (i.e., understood that it was about how low vs. high expertise and low vs. high certainty influence attitudes), or spent less than two minutes on the full questionnaire. With regards to outliers, all the central measures included in this study are bounded scales with up to 9 points, and while we expected variation in participants' responses, we did not expect extreme outliers. Nevertheless, in case of a failed replication, we preregistered to remove univariate outliers with scores of +/− 3 standard deviations or more from the mean, and to report findings both before and after outlier exclusions.

7. Results

7.1. Replication

As part of the direct replication of Study 2 of Karmarkar and Tormala (2010) we analyzed the data using responses from conditions that were included in the original study: low certainty (with reason) vs. high certainty (without reason) for experts vs. non-experts.

Manipulation checks. We tested the responses to the manipulation check questions regarding perceived source expertise and source certainty. First, a 2 × 2 ANOVA with source expertise and source certainty as the independent variables and perceived certainty as dependent variable found support for the main effect of source certainty, $F(1, 506) = 173.13$, $p < .001$, $\eta_p^2 = 0.25$, 95% CI [0.19, 0.31]. Post-hoc comparisons[2] using *t*-tests with Bonferroni correction found that participants reported higher certainty in the high certainty condition ($n = 257$, $M = 7.87$, $SD = 1.24$) than in the low certainty condition ($n = 253$, $M = 5.96$, $SD = 1.96$; $t(426) = 13.00$, $p < .001$, Cohen's $d = 1.17$, 95% CI [0.98, 1.35]). There was no main effect of expertise, nor an expertise × certainty interaction effect, $Fs < 1$.

A 2 × 2 ANOVA with source expertise and source certainty as independent variables and perceived expertise as dependent variable found main effects of source certainty, $F(1, 506) = 5.44$, $p = .020$, $\eta_p^2 = 0.01$, 95% CI [0.00, 0.04], and expertise, $F(1, 506) = 171.48$, $p < .001$, $\eta_p^2 = 0.25$, 95% CI [0.19, 0.31], but no support for an expertise × certainty interaction, $F(1, 506) = 1.26$, $p = .262$, $\eta_p^2 = 0.00$, 95% CI [0.00, 0.02]. Participants in the high expertise condition reported higher perceived expertise ($n = 256$, $M = 7.39$, $SD = 1.51$) than participants in low expertise condition ($n = 254$, $M = 5.24$, $SD = 2.15$; $t(454) = 13.03$, $p < .001$, Cohen's $d = 1.16$, 95% CI [0.97, 1.34]). Additionally, participants rated perceived expertise slightly higher in the high certainty condition ($n = 257$, $M = 6.52$, $SD = 2.06$) than in the low source certainty condition ($n = 253$, $M = 6.12$, $SD = 2.22$; $t(504) = 2.10$, $p = .036$, Cohen's $d = 0.19$, 95% CI [0.01, 0.36]).

Attitudes. The three items measuring attitudes were highly related, $\alpha = 0.93$, and we used the average of the three as an attitude index. See Fig. 2 (top row) for the pattern of responses to the attitude measure across the experimental conditions.[3] A 2 × 2 ANOVA with source expertise and source certainty as the independent variables and attitudes as dependent variable found support for a main effect of source certainty, $F(1, 506) = 4.27$, $p = .039$, $\eta_p^2 = 0.01$, 95% CI [0.00, 0.03]. However, there was no main effect of expertise, $F(1, 506) = 1.35$, $p = .247$, $\eta_p^2 = 0.00$, 95% CI [0.00, 0.02], nor an expertise × certainty interaction effect, $F(1, 506) = 0.84$, $p = .361$, $\eta_p^2 = 0.00$, 95% CI [0.00, 0.02]. The main effect of certainty reflected higher attitude ratings in the high certainty condition ($n = 257$, $M = 7.80$, $SD = 1.10$) than in the low certainty condition ($n = 253$, $M = 7.58$, $SD = 1.26$, Cohen's $d = 0.18$, 95% CI [0.01, 0.36]).

The ANOVA analysis was followed by probing the contrast between certain vs. uncertain recommendation within different levels of expertise on the attitude measure. We did not find support for the original prediction that attitudes will be more favorable after reading reviews from nonexpert sources who expressed certainty ($M = 7.69$, $SD = 1.23$) compared to uncertainty ($M = 7.57$, $SD = 1.26$); $t(252) = 0.77$, $p = .439$, Cohen's $d = 0.10$, 95% CI [−0.15, 0.34]). Contrary to the original prediction that attitudes would be more favorable when expert sources express uncertainty, we found more favorable attitudes when experts expressed certainty ($M = 7.91$, $SD = 0.96$) compared to uncertainty (M

[2] All the post-hoc comparisons following ANOVAs used the same method of Welch's *t*-tests with Bonferroni correction.

[3] For an overview of mean ratings for all dependent variables (manipulation checks, attitude, and involvement ratings) in all the different experimental conditions (both replication and extension conditions) see Tables S9 and S10 in the supplement.

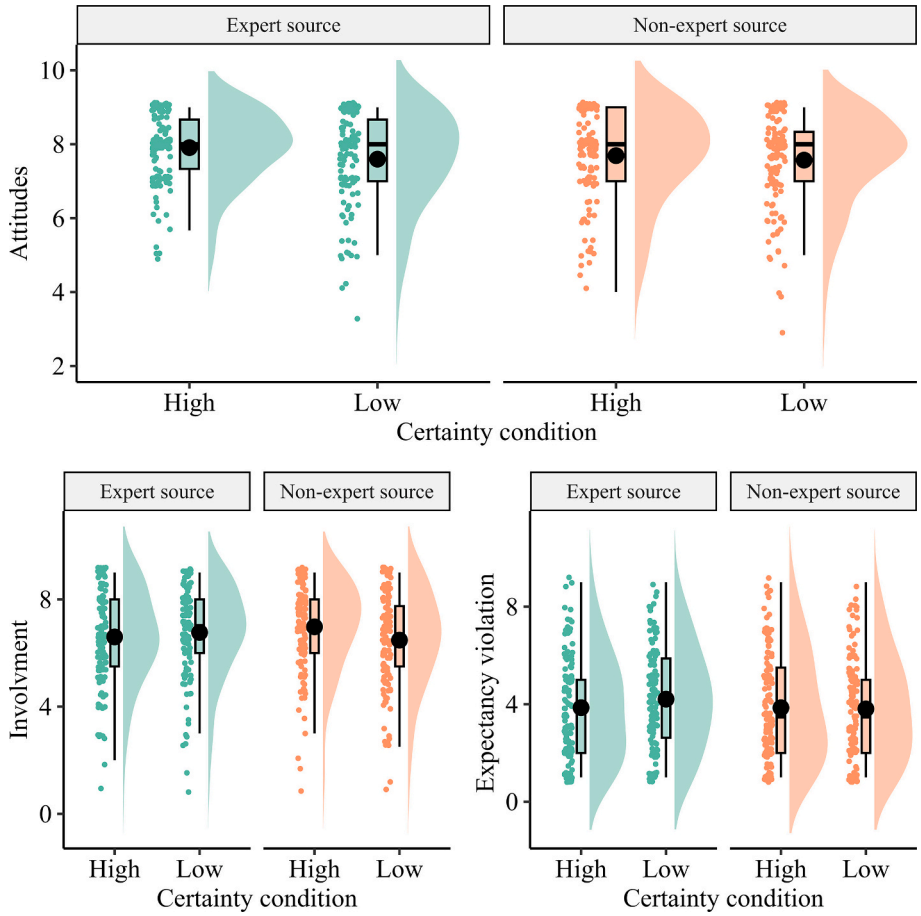

Fig. 2. Distribution of ratings of attitudes (top row), involvement (bottom left), and expectancy violation (bottom right) in experimental conditions directly replicating Karmarkar and Tormala (2010) Experiment 2 (i.e., low certainty with reason vs. high certainty without reason, for non-experts vs. experts).
Note. Colored fields display the distribution of responses. Boxplots display the median, first, and third quartiles. Black circles denote mean values.

= 7.60, *SD* = 1.27); $t(232) = 2.22$, $p = .027$, Cohen's $d = 0.28$, 95% CI [0.03, 0.52].

In summary, we did not find support for Hypothesis 1a, which proposed that non-experts would be more persuasive when expressing certainty and experts would be more persuasive when expressing uncertainty. Instead, the competing Hypothesis 1b, proposing a general positive effect of expressing certainty regardless of expertise, was supported. Table 6 compares the original results with the replication results using LeBel et al.'s (2018, 2019) criteria.

Involvement. The two items measuring involvement were highly related, $\alpha = 0.81$, and an average of the two was used as an involvement index. A 2 × 2 ANOVA with involvement as the dependent variable revealed no main effects for source expertise or source certainty (*F*s < 1.18), but supported an expertise × certainty interaction effect, $F(1, 506) = 4.74$, $p = .030$, $\eta_p^2 = 0.01$, 95% CI[0.00, 0.03]. As illustrated in Fig. 2 (bottom left), participants reported higher involvement with the nonexpert review when the reviewer expressed high ($M = 6.98$, $SD = 1.66$) rather than low certainty ($M = 6.48$, $SD = 1.80$; $t(250) = 2.30$, $p = .023$, Cohen's $d = 0.29$, 95% CI [0.04, 0.53]). In contrast, there was slightly higher involvement with the expert review when the reviewer expressed low ($M = 6.77$, $SD = 1.72$) rather than high certainty ($M = 6.60$, $SD = 1.68$), but this difference was not statistically significant, $t(253) = -0.78$, $p = .437$, Cohen's $d = -0.10$, 95% CI [−0.34, 0.15]). Thus, although these results have similarities to the original study in that an interaction effect is observed, the interaction seems to be driven by reactions to certainty vs. uncertainty in the non-expert condition rather than in the expert condition.

Expectancy violation. We conducted a similar ANOVA analysis with expectancy violation as the dependent variable. The two items measuring expectancy violation were highly related, $\alpha = 0.95$, and an

Table 6
Comparison of replication results with original results, with attitudes as the dependent variable.

	Original statistics	Replication statistics	Original effect size and 95% CI	Replication effect size and 95% CI	Interpretation
Interaction effect, certainty × expertise	$F(1,64) = 9.70$, $p < .004$	$F(1,506) = 0.84$, $p = .361$	$\eta_p^2 = 0.13$ [0.02, 0.29]	$\eta_p^2 = 0.00$ [0.00, 0.02]	No signal – inconsistent
Simple effect of certainty, experts	$F(1,64) = 4.94$, $p < .03$	$F(1, 254) = 4.97$, $p = .027$	$d = -0.56$ [−1.05, −0.05]	$d = 0.28$ [0.03, 0.52]	Signal – inconsistent, opposite direction
Simple effect of certainty, non-experts	$F(1,64) = 4.77$, $p < .04$	$F(1, 252) = 0.60$, $p = .439$	$d = 0.55$ [0.04, 1.04]	$d = 0.10$ [−0.15, 0.34]	No signal - inconsistent

Note. For the simple effects, Cohen's *d* is coded such that a positive sign indicates higher score on the attitude measure in the high certainty than in the low certainty condition, while a negative sign indicates a lower score on the attitude measure in the high vs. the low certainty condition.

average of the two was used as an expectancy violation measure. The results revealed no main effects or an interaction effect (*F*s < 1.15). See Fig. 2 (bottom right) and Table S18 in the supplement for detailed results. These results do not replicate the pattern observed in Karmarkar and Tormala's (2010) Experiment 1, where expert uncertainty (vs. certainty) and non-expert certainty (vs. uncertainty) was found to violate expectations.

7.2. Extensions

7.2.1. Controlling for the reason for (un)certainty

Manipulation check for reason factor. We submitted a 2 × 2 × 2 ANOVA with source expertise, source certainty, and presence or absence of reason as the independent variables and the manipulation check for perceived reason as a dependent variable. There was no support for a main effect of source expertise $F < 1$, but there was support for main effects of certainty ($F(1,1010) = 388.55$, $p < .001$, $\eta^2_p = 0.28$, 95% CI [0.23, 0.32]) and reason ($F(1, 1010) = 32.76$, $p < .001$, $\eta^2_p = 0.03$, 95% CI [0.01, 0.06]) along with a certainty × reason interaction, ($F(1, 1010) = 16.00$, $p < .001$, $\eta^2_p = 0.02$, 95% CI [0.00, 0.03]). Most importantly, the results suggest the reason manipulation worked as intended: participants in conditions where a reason for certainty or uncertainty was provided perceived there to be a better reason ($n = 507$, $M = 6.70$, $SD = 2.11$) than participants in conditions where no reason was provided ($n = 511$, $M = 6.03$, $SD = 2.34$; $t(1007) = 4.80$, $p < .001$, Cohen's $d = 0.30$, 95% CI [0.18, 0.42]).[4]

7.2.2. Robustness of original findings

As part of testing the robustness of the original findings while accounting for the role of reason for (un)certainty, we conducted a 2 (reason: yes vs. no) × 2 (source expertise: expert vs. non-expert) × 2 (source certainty: certain vs. uncertain) ANOVA with attitudes as the dependent variable. See Fig. 3 for the pattern of responses to the attitude measure across the experimental conditions. The analysis involved reason, source expertise, source certainty, and the expertise × certainty interaction as predictors of the attitude measure. The results revealed a main effect of certainty, $F(1,1013) = 15.32$, $p < .001$, $\eta^2_p = 0.01$, 95% CI [0.00, 0.03]. Besides this there was no support for other main or interaction effects (*F*s < 1.45). The main effect of certainty reflected more favorable attitudes when the reviewer expressed certainty ($M = 7.84$, $SD = 1.08$) compared to uncertainty ($M = 7.56$, $SD = 1.21$); $t(1001) = 3.92$, $p < .001$, Cohen's $d = 0.25$, 95% CI [0.12, 0.37]). In summary, these findings were similar to the results in the conditions directly replicating the original study: they support *H1b* (confidence heuristic; certainty as a direct predictor) rather than *H1a* (expertise × certainty interaction term, $F < 1$). Furthermore, the findings did not support Hypothesis 2, that reason would be a direct predictor of attitudes ($F < 1$).

We conducted a similar set of analyses with involvement and expectancy violation as the outcome variables. The detailed results are reported in the supplementary materials sections (see Tables S23 – S28). When the outcome variable was expectancy violation, we found a main effect of certainty ($F(1, 1013) = 12.04$, $p = .001$, $\eta^2_p = 0.01$, 95% CI [0.00, 0.03]), with the review being rated lower on the expectancy violation measure in the high certainty condition ($M = 3.71$, $SD = 2.12$) than in the low certainty condition ($M = 4.16$, $SD = 2.06$; $t(1015) = -3.46$, $p < .001$, Cohen's $d = -0.22$, 95% CI [−0.34, −0.09]). We also found support for a main effect of reason ($F(1, 1013) = 5.38$, $p = .021$, $\eta^2_p = 0.01$, 95% CI [0.00, 0.02]), with the review being rated lower on the expectancy violation measure in the presence of a reason ($M = 3.78$, $SD = 2.09$) than without one ($M = 4.09$, $SD = 2.10$; $t(1016) = -2.31$, $p = .021$, Cohen's $d = -0.14$, 95% CI [−0.27, −0.02]). Besides this the results revealed no support for other main effects or the interaction effect (*F*s < 1.68). When the outcome variable was involvement there was no support for any of the predictors (*F*s < 1.89). It is important to note here that in Karmarkar and Tormala (2010), the incongruity effect was hypothesized to arise as an effect of expectancy violation, with expert uncertainty and novice certainty violating expectations and increasing involvement with the message. Overall, the results for involvement and expectancy violation further demonstrate that our current results do not support the original incongruity hypothesis.

7.2.3. Need for closure as an individual level moderator

As an extension, we also investigated need for closure (NFC) as a potential individual level moderator. We averaged the responses to the 15-item NFC scale ($\alpha = 0.90$). To test the moderation effect of need for closure, we conducted a linear regression analysis including expertise, reason, certainty, NFC, and two interaction terms (expertise × certainty and certainty × NFC) as predictors of the attitude measure. The analysis showed a main effect of NFC but failed to find an interaction effect (*H3*) between certainty and NFC (see Table 7). The main effect reflected that NFC was positively associated with attitudes, such that individuals with higher NFC reported more positive attitudes in response to the positive review they had read. The results were similar for analyses with involvement and expectancy violation as dependent variables (see Table S29-S31). In other words, we found no support for *H3*.

7.2.4. Exploratory analyses

The participants reported on four types of perceived uncertainty: epistemic, aleatory, external, and internal uncertainty. We conducted a series of ANOVAs to investigate the interactions among three factors: reason (yes vs. no), source expertise (expert vs. non-expert), and source certainty (certain vs. uncertain), with each type of perceived uncertainty treated as a separate dependent variable. For all four types of uncertainty, there was a main effect of certainty, indicating that perceived type of uncertainty depended on whether the reviewer was certain or uncertain. There were no main effects of expertise, no expertise × reason or expertise × reason × certainty interaction for any measure. Some other effects appeared for only some measures (e.g., main effect of reason and reason × certainty interaction for internal and external uncertainty, see Table S33-S36). Additionally, we conducted analyses that accounted for type of uncertainty as a predictor of attitudes (see Table S37-S41). Here, we found that perceived degree of aleatory uncertainty negatively predicted attitudes, perceived degree of epistemic and external uncertainty positively predicted attitudes, while degree of internal uncertainty did not predict attitudes by itself.

7.2.5. Results after exclusions

Since our results indicated a failed replication, we also reran the analyses after removing participants based on our exclusion criteria and outlying responses, i.e., participants scoring +/− 3 standard deviations or more from the mean on the attitude measure. This led to one changed result in conditions directly replicating Karmarkar and Tormala's Study 2. Specifically, a 2 × 2 ANOVA with source expertise and source certainty as the independent variables and attitudes as dependent variable showed no statistically significant main effect of certainty, $F(1, 494) = 2.41$, $p = .121$, $\eta^2_p = 0.00$, 95% CI [0.00, 0.02], no main effect of expertise, $F(1, 494) = 0.73$, $p = .395$, $\eta^2_p = 0.00$, 95% CI [0.00, 0.02], nor an expertise × certainty interaction effect, $F(1, 494) = 1.46$, $p = .228$, $\eta^2_p = 0.00$, 95% CI [0.00, 0.02]. However, when all conditions were included, we did find support for a main effect of certainty, $F(1, 992) = 13.16$, $p < .001$, $\eta^2_p = 0.01$, 95% CI [0.00, 0.03], while no other main or interaction effects were statistically significant, all *F*s < 1.75, all *p*s > 0.18. See Table S42 and S43 for detailed results.

[4] The main effect of certainty reflected a better perceived reason for certainty than for uncertainty. The interaction showed that the reason manipulation had a stronger effect in the low certainty condition, $d = 0.51$, 95% CI [0.33, 0.69], than in the high certainty condition, $d = 0.14$, 95% CI [−0.03, 0.32].

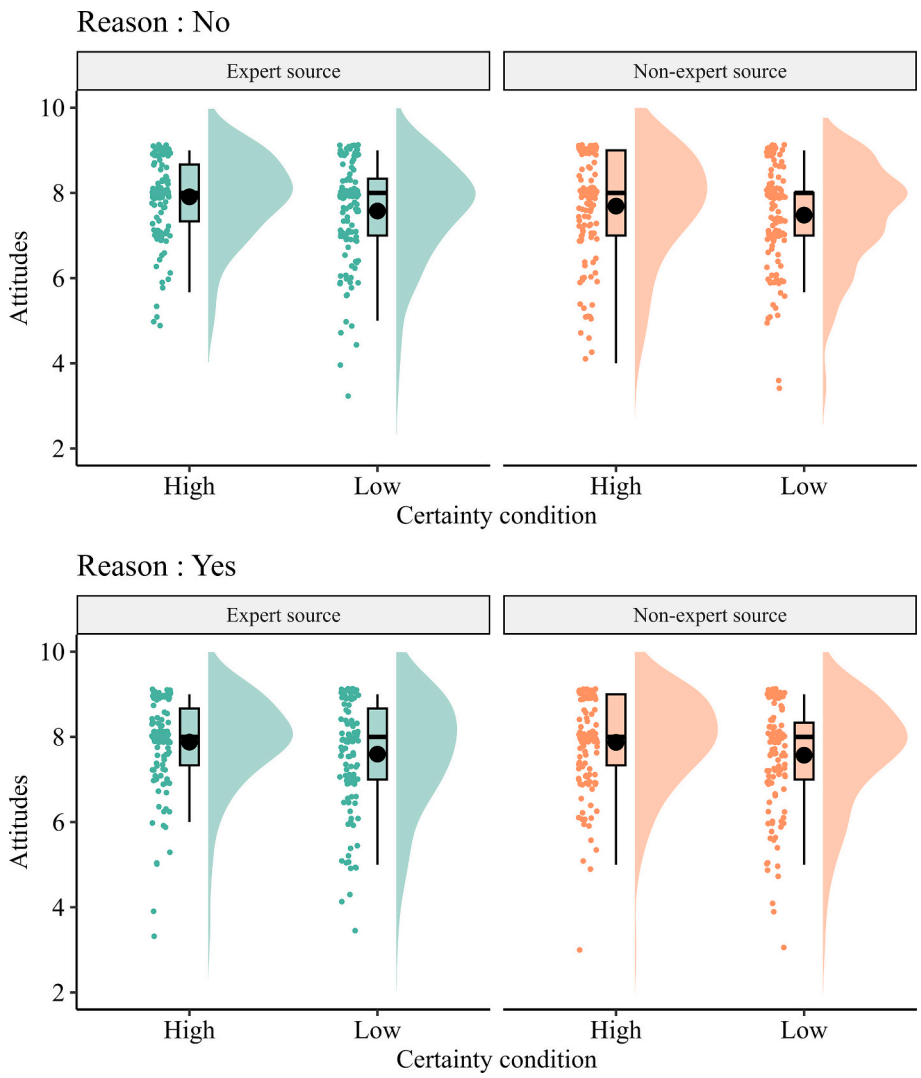

Fig. 3. Distribution of attitude ratings by certainty, expertise, and reason experimental conditions.
Note. Colored fields display the distribution of responses. Boxplots display the median, first, and third quartiles. Black circles denote mean values.

Table 7
Results of linear regression analysis with attitudes as outcome measure.

	Attitudes		
Predictors	*Estimates*	*95% CI*	*p*
Intercept	7.27	6.94–7.60	**<0.001**
Expertise (Low vs. High)	−0.06	−0.16–0.04	0.214
Reason (No vs. Yes)	−0.05	−0.15–0.05	0.325
Certainty (Low vs. High)	−0.15	−0.61–0.32	0.531
Need for Closure (NFC)	0.11	0.03–0.19	**0.010**
Certainty x Expertise	0.03	−0.11–0.17	0.645
Certainty x NFC	−0.01	−0.13–0.10	0.813
Observations	1017		
R^2 / R^2 adjusted	0.024 / 0.018		

Note. Estimates are non-standardized.

8. Discussion

The question of how to communicate uncertainty while retaining credibility seems increasingly important in an unpredictable world. The intriguing findings of Karmarkar and Tormala (2010) suggested that experts and non-experts could use different strategies: while non-experts were more persuasive when they were certain, people were more persuaded by experts who expressed their opinions with uncertainty. We set out to replicate Karmarkar and Tormala's Study 2. Our findings, however, diverged from the original study in several key aspects.

Most importantly, in a well-powered direct replication, we did not find support for the interaction between source expertise and certainty proposed in the original study's *incongruity hypothesis (H1a)*. Instead, we found evidence supporting the competing *confidence heuristic hypothesis (H1b)*: both experts and non-experts were more persuasive when expressing their opinions with certainty rather than uncertainty. Thus, for experts the effect of expressing uncertainty was in the opposite direction of the effect observed by Karmarkar and Tormala (2010), and in fact, the positive effect of certainty on attitudes was stronger for experts than for non-experts.

In the conditions directly replicating the original, we did observe an interaction effect of source certainty and expertise on involvement. Specifically, high certainty from non-experts significantly enhanced involvement with the review, while there was a non-significant difference in the opposite direction for experts. It is noteworthy that the significant interaction was driven by the increased involvement with high certainty non-expert reviews, and not by increased involvement with uncertain expert reviews. The increased persuasion and involvement due to expert uncertainty is the most surprising and counterintuitive finding in the original study, so even though this interaction effect is similar to what Karmarkar and Tormala (2010) found, it cannot be said

to give strong support to the incongruity hypothesis. Furthermore, there was no similar pattern when examining expectancy violation, making the overall results more in line with the idea that certainty generally is persuasive, regardless of source expertise. The results for involvement and expectancy violation mean that we did not replicate the proposed mechanisms behind the incongruity effect.

We also extended the original study in several ways. First, we investigated the reason for (un)certainty as a potential confound. The manipulation check showed that participants perceived a better reason for (un)certainty in conditions where a reason was provided, indicating that our manipulation functioned as intended. However, contrary to Hypothesis 2, providing a reason for (un)certainty did not significantly predict attitudes, and reason did not interact with any other independent variables. These results suggest that the reason given for uncertainty in the original study was not a strong confound.

Additionally, we also included need for closure (NFC) as a potentially relevant individual difference. We hypothesized that those with higher NFC would be more positive toward certain reviews and less positive toward uncertain reviews, but found no support for this prediction (*H3*). Instead, NFC had a main effect on attitudes, with higher scores on NFC associated with more positive attitudes. A plausible post-hoc explanation of this result is that high NFC individuals, after reading the positive review, did not feel the need to think more deeply about the matter, but simply incorporated their initial positive impression into their evaluation of the restaurant. Given that NFC is an individual's desire for a firm view and an aversion toward ambiguity, participants high on NFC may have anchored their evaluation on the rating ("4 out of 5") which was presented on a numeric scale, with less regard for source expertise and confidence. This interpretation is in line with the tendencies to "seize and freeze" that are part of need for closure (Kruglanski & Webster, 1996). Future work could explore the role of NFC with information related to expertise and confidence presented as scale ratings.

Further exploratory analyses included measures of the type of uncertainty. These analyses showed that perceptions of type of uncertainty (1) can be influenced by the manipulations and (2) can predict attitudes. However, as perceived type of uncertainty was influenced by the level of certainty, it is not straightforward to draw strong conclusions from these results. We still believe this could be an important factor to understand different effects of communicating uncertainty (Løhre & Teigen, 2023; Ülkümen et al., 2016), and hope to pursue this in future research.

While we did not find support for our extension hypotheses, the full sample gave further support for the main effect of certainty on attitudes. Even when controlling for the reason for (un)certainty, participants were more persuaded by reviews where reviewers expressed high certainty. In fact, in the full sample, the effect of certainty was statistically significant for both experts ($p = .002$, $d = 0.28$) and non-experts ($p = .015$, $d = 0.22$). Results after excluding participants based on exclusion criteria added a small wrinkle, as the main effect of certainty was not statistically significant in conditions directly replicating the original study. However, when all conditions were included, the main effect of certainty was supported even after exclusions. Overall, we must conclude that our study failed to replicate the findings in the original, and that we find support for the confidence heuristic hypothesis (*H1b*) rather than the incongruity hypothesis (*H1a*).

An important question is why the results in the original and this replication diverge. We can first exclude the possibility that our study failed in effectively manipulating the independent variables. Participants clearly distinguished between reviews with low and high certainty, and between expert and non-expert reviewers, as evidenced by the large effect sizes observed for the manipulation checks ($ds > 1$). Even the new manipulation of reason for (un)certainty had the intended effect, albeit with a small effect size ($d = 0.30$). These results also indicate that most participants paid attention, and make it implausible that other results represent noise or random answers.

The replication differed from the original study in some minor details. We used online participants (vs. students in the lab), the review was said to stem from a website collecting reviews rather than from an online journal, and prices were updated to current standards. We deem it unlikely that these minor differences could explain the discrepant results. A more likely candidate explanation is contextual differences. Replication studies are often criticized for not taking contextual sensitivity into account (Nosek et al., 2022), and we are open to the possibility that context could matter here. To name some potentially relevant events since the publication of Karmarkar and Tormala in 2010: online reviews from both experts and non-experts have arguably become more common; we have been exposed to leaders like Donald Trump who use excessive certainty and confidence to persuade (Moore & Bazerman, 2022); the COVID-19 pandemic (Kerr et al., 2023) and later the war in Ukraine have involved experts (and non-experts) communicating about highly uncertain topics. Events like these could influence associations to uncertainty, experts, or both.

One speculation we could make in this regard concerns the two variables involvement and expectancy violation. The effect proposed by Karmarkar and Tormala was hypothesized to stem from the incongruency between expertise and certainty as persuasion variables, with people finding it surprising (in violation of expectations) when novices express certainty and experts express uncertainty. This in turn should lead to higher involvement with the attitudinal message and increased persuasion. The participants in the current replication apparently did not find expert uncertainty or novice certainty particularly surprising. Perhaps they have gotten used to (over)confident user reviews in online settings, and to experts expressing uncertainty in different settings. This possibility could be explored in future studies which establish the common associations in the relevant persuasion domain, with the expectation that the original finding would replicate when expert certainty and novice uncertainty is the norm.

Regardless of what explains the different results in the original and this replication, we believe the most reasonable conclusion from the present findings is that there is more support for the confidence heuristic hypothesis, which is already well-established in the literature, than for the incongruity hypothesis. Relatedly, the current findings may have implications for the recent theoretical framework proposed by Hussein and Tormala (2021), where conveying uncertainty is seen as an "act of receptiveness" that could increase persuasion, especially for expert sources. While it seems reasonable that openness about uncertainty can be more persuasive than certainty at least in some cases, the current results do not provide support for this theoretical proposition, but rather suggest that both experts and non-experts stand to benefit from expressing their opinions with certainty. However, the effect size is rather small, so an alternative framing of these results could be that there is not much to lose from being open about uncertainty. This resonates with recent findings indicating that trust in statistics and expert sources is not necessarily much reduced by communicating uncertainty (Kerr et al., 2023; van der Bles et al., 2020).

Both the original study and the current replication are limited by the use of a consumer context, specifically a restaurant review. We started our paper by discussing the far-reaching implications of a potential persuasive effect of uncertainty for experts involved in a variety of highly important topics. The current results cannot necessarily inform experts involved in such topics about whether or how they should communicate uncertainty. We believe it is crucial to continue investigating this question in a variety of settings and contexts, both in the laboratory and outside.

Financial disclosure/funding

The project is supported by an internal research grant from BI Norwegian Business School.

CRediT authorship contribution statement

Erik Løhre: Writing – review & editing, Writing – original draft,

Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Subramanya Prasad Chandrashekar:** Writing – review & editing, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Lewend Mayiwar:** Writing – review & editing, Visualization, Validation, Methodology, Funding acquisition, Conceptualization. **Thorvald Hærem:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Data availability

All study materials, code, and collected data are provided on: https://osf.io/hbjyv/

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2024.104619.

References

Ahluwalia, S. C., Edelen, M. O., Qureshi, N., & Etchegaray, J. M. (2021). Trust in experts, not trust in national leadership, leads to greater uptake of recommended actions during the COVID-19 pandemic. *Risk, Hazards & Crisis in Public Policy, 12*(3), 283–302. https://doi.org/10.1002/rhc3.12219

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., … Nosek, B. A. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality, 27*(2), 108–119.

van der Bles, A. M., van der Linden, S., Freeman, A. L. J., & Spiegelhalter, D. J. (2020). The effects of communicating uncertainty on public trust in facts and numbers. *Proceedings of the National Academy of Sciences, 117*(14), 7672–7683. https://doi.org/10.1073/pnas.1913678117

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. https://doi.org/10.1016/j.obhdp.2006.07.001

Brewer, N., & Burke, A. (2002). Effects of testimonial inconsistencies and eyewitness confidence on mock-juror judgments. *Law and Human Behavior, 26*(3), 353–364. https://doi.org/10.1023/A:1015380522722

Bühren, C., Meier, F., & Pleßner, M. (2023). Ambiguity aversion: Bibliometric analysis and literature review of the last 60 years. *Management Review Quarterly, 73*(2), 495–525. https://doi.org/10.1007/s11301-021-00250-9

Corner, A., Lewandowsky, S., Phillips, M., & Roberts, O. (2015). *The uncertainty handbook* (p. 18). University of Bristol.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics, 75*(4), 643–669. https://doi.org/10.2307/1884324

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fox, C. R., & Ülkümen, G. (2017). Comment on Løhre & Teigen (2016). "There is a 60% probability, but I am 70% certain: Communicative consequences of external and internal expressions of uncertainty". *Thinking & Reasoning, 23*(4), 483–491. https://doi.org/10.1080/13546783.2017.1314939

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

Grant, A. (2021). *Think again: The power of knowing what you don't know*. Penguin.

Gustafson, A., & Rice, R. E. (2020). A review of the effects of uncertainty in public science communication. *Public Understanding of Science, 29*(6), 614–633. https://doi.org/10.1177/0963662520942122

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes, 70*(2), 117–133. https://doi.org/10.1006/obhd.1997.2697

Hussein, M. A., & Tormala, Z. L. (2021). Undermining your case to enhance your impact: A framework for understanding the effects of acts of receptiveness in persuasion. *Personality and Social Psychology Review, 25*(3), 229–250. https://doi.org/10.1177/10888683211001269

Isager, P. M., Van Aert, R. C. M., Bahník, Š., Brandt, M. J., DeSoto, K. A., Giner-Sorolla, R., … Lakens, D. (2023). Deciding what to replicate: A decision model for replication study selection under resource and knowledge constraints. *Psychological Methods, 28*(2), 438–451. https://doi.org/10.1037/met0000438

Juanchich, M., Gourdon-Kanhukamwe, A., & Sirota, M. (2017). "I am uncertain" vs "It is uncertain". How linguistic markers of the uncertainty source affect uncertainty communication. *Judgment and Decision Making, 12*(5), 445–465.

Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition, 11*(2), 143–157. https://doi.org/10.1016/0010-0277(82)90023-3

Kardes, F. R., Fennis, B. M., Hirt, E. R., Tormala, Z. L., & Bullington, B. (2007). The role of the need for cognitive closure in the effectiveness of the disrupt-then-reframe influence technique. *Journal of Consumer Research, 34*(3), 377–385. https://doi.org/10.1086/518541

Karmarkar, U. R., & Tormala, Z. L. (2010). Believe me, I have no idea what I'm talking about: The effects of source certainty on consumer involvement and persuasion. *Journal of Consumer Research, 36*(6), 1033–1049. https://doi.org/10.1086/648381

Kause, A., Bruine de Bruin, W., Domingos, S., Mittal, N., Lowe, J., & Fung, F. (2021). Communications about uncertainty in scientific climate-related findings: A qualitative systematic review. *Environmental Research Letters, 16*(5), Article 053005. https://doi.org/10.1088/1748-9326/abb265

Kelp, N. C., Witt, J. K., & Sivakumar, G. (2022). To vaccinate or not? The role played by uncertainty communication on public understanding and behavior regarding COVID-19. *Science Communication, 44*(2), 223–239. https://doi.org/10.1177/10755470211063628

Kerr, J., Van Der Bles, A.-M., Dryhurst, S., Schneider, C. R., Chopurian, V., Freeman, A. L. J., & Van Der Linden, S. (2023). The effects of communicating uncertainty around statistics, on public trust. *Royal Society Open Science, 10*(11), Article 230604. https://doi.org/10.1098/rsos.230604

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review, 103*(2), 263–283. https://doi.org/10.1037/0033-295X.103.2.263

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science, 4*(1). https://doi.org/10.1177/2515245920951503, 251524592095150.

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science, 1*(3), 389–402. https://doi.org/10.1177/2515245918787489

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology, 3*, 9.

Løhre, E. (2018). Stronger, sooner, and more certain climate change: A link between certainty and outcome strength in revised forecasts. *Quarterly Journal of Experimental Psychology, 71*(12), 2531–2547. https://doi.org/10.1177/1747021817746062

Løhre, E., & Teigen, K. H. (2016). There is a 60% probability, but I am 70% certain: Communicative consequences of external and internal expressions of uncertainty. *Thinking & Reasoning, 22*(4), 369–396. https://doi.org/10.1080/13546783.2015.1069758

Løhre, E., & Teigen, K. H. (2023). When leaders disclose uncertainty: Effects of expressing internal and external uncertainty about a decision. *The Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218231204350

Lüdecke, D., Makowski, D., Ben-Shachar, M. S., Patil, I., & Wiernik, B. M. (2022). Easystats: Framework for easy statistical modeling, visualization, and reporting (0.5.2) [computer software]. https://CRAN.R-project.org/package=easystats.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology, 61*(1), 13.

Moore, D. A., & Bazerman, M. H. (2022). Leadership & overconfidence. *Behavioral Science & Policy, 8*(2), 59–69. https://doi.org/10.1177/237946152200800206

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., … Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology, 73*, 719–748. https://doi.org/10.1146/annurev-psych-020821- 114157

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Otto, A. S., Clarkson, J. J., & Kardes, F. R. (2016). Decision sidestepping: How the motivation for closure prompts individuals to bypass decision making. *Journal of Personality and Social Psychology, 111*(1), 1–16. https://doi.org/10.1037/pspa0000057

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Communication and persuasion* (pp. 1–24). Springer.

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. (2010). 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

Roets, A., & Van Hiel, A. (2011). Item selection and validation of a brief, 15-item version of the need for closure scale. *Personality and Individual Differences, 50*(1), 90–94. https://doi.org/10.1016/j.paid.2010.09.004

Smithson, M. (2008). Psychology's ambivalent view of uncertainty. In *Uncertainty and risk—Multidisciplinary perspectives*. Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781849773607-26/psychology-ambivalent-view-uncertainty-michael-smithson.

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. https://doi.org/10.1006/obhd.2000.2926

Teigen, K. H. (1988). The language of uncertainty. *Acta Psychologica, 68*, 27–38. https://doi.org/10.1016/0001-6918(88)90043-1

Teigen, K. H., & Løhre, E. (2017). Expressing (un)certainty in no uncertain terms: Reply to Fox and Ülkümen. *Thinking & Reasoning, 23*(4), 492–496. https://doi.org/10.1080/13546783.2017.1314965

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. https://doi.org/10.1111/j.1467-9280.2007.01847.x

Tormala, Z. L., & Rucker, D. D. (2007). Attitude certainty: A review of past findings and emerging perspectives. *Social and Personality Psychology Compass, 1*(1), 469–492. https://doi.org/10.1111/j.1751-9004.2007.00025.x

Ülkümen, G., Fox, C. R., & Malle, B. F. (2016). Two dimensions of subjective uncertainty: Clues from natural language. *Journal of Experimental Psychology: General, 145*(10), 1280–1297. https://doi.org/10.1037/xge0000202

Walters, D. J., Ülkümen, G., Tannenbaum, D., Erner, C., & Fox, C. R. (2023). Investor behavior under epistemic vs. aleatory uncertainty. *Management Science, 69*(5), 2761–2777. https://doi.org/10.1287/mnsc.2022.4489

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*(6), 1049–1062. https://doi.org/10.1037/0022-3514.67.6.1049

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences, 41*, Article e120. https://doi.org/10.1017/S0140525X17001972