Research Repository

Do claims about certainty make estimates less certain?

Accepted for publication in Cognition.

Do claims about certainty make estimates less certain?

Karl Halvor Teigen[1] Marie Juanchich[2]

[1]University of Oslo, Norway*

[2]University of Essex, UK

*Corresponding author:

Karl Halvor Teigen

Department of Psychology

University of Oslo

0317 Oslo, Norway

k.h.teigen@psykologi.uio.no

Co-author:

Marie Juanchich

Department of Psychology

University of Essex, UK

m.juanchich@essex.ac.uk

Abstract

Predictions and estimates are sometimes qualified as *certain.* This epistemic marker occupies a privileged position at the top of scales of verbal probability expressions, reflecting probabilities close to 1. But such statements have rarely been compared to plain, unqualified statements in which certainty is not mentioned. We examined in nine studies (N = 2,784) whether statements explicitly claimed to be certain are perceived as (1) more (or less) credible, (2) more (or less) precise, and (3) more (or less) strongly based upon evidence, compared to plain, unmarked declarative statements. We find, in apparent contrast with assumptions made by the standard scales, that "certain" are often judged to be less trustworthy, less reliable, and held with lower confidence than unmarked statements. Plain, declarative statements are further assumed to be more precise, while certainty implies that more extreme outcomes are possible. When it is certain that Henry made four errors, it is clear he did not commit less than four, but he might have committed five errors or more. Thus certainty can indicate lower bounds of an interval whose upper bounds are not defined, and certainty statements are consequently more ambiguous than estimates that do not mention certainty. At least-interpretations of certainty affect the interpretation of options in risky choice problems, where "200 lives will be saved" was deemed by a majority to mean exactly 200, while "it is certain that 200 will be saved", could mean 200-600 lives. We also find that credibility is affected by type of certainty, with impersonal certainty ("it is certain") perceived to be more accurate than personal certainty ("I am certain"). Moreover, certainty can reveal that that the speaker's estimate is based on subjective judgments and guesswork rather than upon objective evidence. These findings have implications for communicating estimates. Climate predictions are believed to be more consensual when claims of certainty are omitted. To convey certainty it may be better not to mention that one is certain.

Key words

Certainty verbal probabilities framing pragmatics credibility preciseness

Highlights

- Statements claimed to be *certain* are compared to plain, unmarked statements
- Unmarked estimates are often judged to be more believable and less contentious
- Estimates claimed to be certain are often given an *at least*-interpretation
- Plain, unmarked estimates are supposed to be more strongly based on evidence
- "I am certain" is less credible than "It is certain"

1. Introduction

Degrees of certainty and uncertainty are integral parts of our lives, from daily routines (e.g., I think her birthday is in March), to information more critical to individual's as well as government's decision-making (e.g., it is possible the sea will rise 1 meter). In language uncertainty is conveyed in many ways, collectively known as verbal probabilities (words and phrases like "uncertain", "possible", and "very likely"). For recent overviews of the meaning of such phrases, see Collins and Hahn (2018) and Dhami and Mandel (2022).

Certainty is a precious commodity. Lack of certainty might be taken as a cue that the speaker lacks competence. For example, communicating verbal uncertainty about an estimate (vs. a factual statement) was shown to decrease trust in the suggested values (but not trust in the source of the estimates; van der Bles et al., 2020). Hence, leaders might be reluctant to communicate uncertainty (Løhre & Teigen, 2023). The confidence heuristics posit that the confidence of a speaker is used as a proxy for their competence (Price & Stone, 2004) and has received consistent empirical support (for a recent replication, see Løhre et al., 2024). Even brands leverage the expression of certainty in their advertisements to increase engagement on social media (Pezzuti et al., 2021). Expressing certainty is portrayed as a sign of strength, competence and trustworthiness, in a so called post truth era, when it is often difficult to distinguish between truthful and false information (Lewandowsky, 2020).

But certainty can be communicated in several ways. One is to state, explicitly, that a claim is certain. Another is to state the facts plainly without additional comments about their certainty. A climate expert can say "it is certain that the sea level will rise" or simply "the sea level will rise". We argue in this paper that explicit claims of certainty, contrary to their intentions, do not always serve to make assertions more persuasive or compelling compared to plain, unqualified statements. To the best of our knowledge, a comparison of these two kinds of statements has previously not been attempted.

1.1. A difference in probability?

In lists of verbal probability expressions *certain* holds a privileged position at the top (e.g., Clark, 1990; Clarke et al., 1992; Hamm, 1991; Mandel & Irwin, 2021; Reagan et al., 1989; Willems et al., 2020; Witteman & Renooij, 2003). This term has been assumed to convey a 100% probability; in fact when people are asked to assess subjective probabilities on numeric scales from 0 to 100, the top end of this scale is typically labelled *certain.* Lexicons developed for expressing probabilities underscore this notion by recommending probabilities close to 1 as *virtually certain* (Mastrandrea et al. , 2010) or as *almost certain* (European Food Association et al., 2018; Mandel & Irwin, 2021; Wintle et al., 2019). So, when speakers say that an event is *certain*, the statement comes with an assurance that its probability is extremely high. Without this guarantee, recipients must judge for themselves. We might therefore assume that statements qualified as “certain” would appear as more credible than unqualified assertions, which in principle could range from mere opinions to established facts.

However, explicit claims of certainty may be considered unnecessary because of the “truth bias” of recipients: Plain, unqualified statements will in most cases be accepted unconditionally, as speakers are assumed by default to tell the truth (Levine, 2014), and listeners find it easier to believe than doubt what they are told (Gilbert, 1991). Speakers who describe an obvious and indisputable feature of a situation, would normally not bother to add assurances about its certainty. One does not say: “I am certain it is raining”, unless this is a matter of debate. It follows that explicit certainty can be less, not more believable, on two accounts. (i) They may alert recipients to the information being potentially contentious (Isberner et al., 2013). From a conversational perspective (Grice, 1975), statements said to be certain may presuppose an actual or imagined conversation partner that holds a divergent view. Alternatively, as implied by the epistemic vigilance hypothesis (Sperber et al., 2010), utterances may require an explicit marker of veracity if they are coming from a source that cannot be trusted unconditionally. Factual statements coming from a reliable source (or have less relevance for the receiver) are not in need of having their certainty declared (Gilbert, 1991; Levine, 20149. (ii) Moreover statements about certainty imply that a *human judgment* has been performed, over and beyond a mere description of objective facts, especially if we assume that comprehension of an utterance include both a comprehension of its content and of the precondition for its acceptance. For this reason, certainty and other verbal probabilities have been called epistemic qualifiers. There might accordingly be occasions when the presence of an epistemic qualifier, even one that posits confidence and conviction, is *less* compelling than its absence.

Thus, with respect to probability perception, arguments exist for two alternative hypotheses:

H1a: Statements qualified by “certain” might convey a *higher* probability than just plain unqualified statements. This follows from a conventional interpretation of what “certain” means.

H1b: Statements qualified by “certain” are more questionable and dependent on judgmental processes than just plain statements and may accordingly convey a *lower* probability. This is a novel prediction to be examined in the present studies. We expected to find results both in support of H1a and H1b, depending on the speaker and the topic (what is claimed to be certain).

1.2. A difference in precision?

A second issue to be considered when comparing statements with or without a certainty qualifier is whether claims about certainty convey a message *beyond* their probabilistic meaning, namely in terms of exactitude.

Again, two divergent hypotheses could be formulated.

H2a: Statements qualified by "certain" convey a *more* precise and reliable estimate than just plain factual statements.

H2b: Statements qualified by "certain" convey a *less* precise and reliable estimate than just plain statements.

The first of these two hypotheses seems to follow from the "confidence heuristic" (Price & Stone, 2004), where speakers' self-proclaimed confidence is taken as a cue for competence. Statements qualified by certainty can be supposed to indicate superior knowledge and hence be perceived as more exact. People believe that speakers using precise numbers are more confident than speakers using round numbers (Jerez-Fernandez et al., 2014), and often assume that narrow interval predictions convey more certainty than wide ones (Løhre & Teigen, 2017: Løhre et al., 2019), despite the fact that those narrow estimates are more risky and can more easily be wrong. Claims of certainty might correspondingly entail assumptions about expertise and precision.

But certainty statements are sometimes used to indicate the lower bound of a prediction interval in a distribution where no single outcome is 100% likely to occur. In studies of how "certain" and the modal "will" were used to predict quantities in such distributions, people selected the smallest value of a distribution of potential values (Teigen & Filkuková, 2013; Teigen et al., 2014), indicating that they included all outcomes from the stated value and upwards, that together would be 100% certain to occur. A "certain" estimate might accordingly indicate a value that could "at least" occur, the lower bound of an open range rather than a point prediction (Juanchich et al., 2013). This would make certainty statements more approximate and imprecise than unqualified estimates.

1.3. Two kinds of uncertainty and certainty

The word "certain" may suggest that the estimate is based on human judgment rather than upon measurements or objective facts, as it can be argued that certainty (like uncertainty) primarily describes an epistemic state (Fox & Ülkümen, 2017; Ülkümen et al., 2016). Therefore we could expect plain statements to be perceived as based more strongly on objective evidence, whereas certainty-statements would be perceived as judgment-based. However, the source of certainty or uncertainty can still be construed as attributable to internal (subjective) or to external (objective) factors. This leads us to a third set of hypotheses, namely whether the interpretation of certainty depends on whether it refers to an internal (subjective) state of knowledge or to external (objective) facts.

In theories of probability, there is a long tradition to distinguish between probabilities of an epistemic and an aleatory kind (Gillies, 2000; Hacking, 1975; Kahneman & Tversky, 1982). *Epistemic* probabilities are assumed to reflect people's state of knowledge, as when judges indicate that a suspect's guilt is "beyond reasonable doubt", whereas *aleatory* probabilities are attributed to external features of the situation, for instance amount of variability and randomness, as when a weather forecaster predicts there is a 90% chance of snow next week. While numerical probabilities are more often used to reflect aleatory uncertainty, verbal probabilities can reflect both [1]. Within verbal probabilities, the epistemic-aleatory distinction can be reflected in the quantifiers used. Ülkümen et al. (2016) suggested that words like "confidence" and "uncertainty" reflect probabilities of an internal, epistemic kind, whereas "likelihood" and "chance" are better suited for describing external, aleatory probabilities.

However, some verbal terms, like "certain", are equivocal and can be conceived as both internal or external depending on their textual setting, as governed by the grammatical subject. Personal phrases like "I am X% certain" describe the speaker's epistemic state, whereas impersonal phrases, like "it is X% certain" are presumably based more strongly on external, impersonal frequencies or trends (Juanchich et al., 2017; Teigen, 2022; Teigen & Løhre, 2017). One study comparing such expressions showed that people ascribed higher internal than external certainty to the same event (Løhre & Teigen, 2016). For instance, when asked to predict the winner of a sports event they said "I am 70% certain", but in a parallel, impersonal context they claimed more modestly that "it is 60% certain". In contrast to these findings, Fox et al. (2011) suggested that listeners place more weight on personal than on impersonal certainty, presumably because of the responsibilities involved, while Juanchich et al. (2017) observed that this applied mainly to expert statements. Statements by non-experts were in their study regarded more trustworthy in an impersonal than in a personal format.

The third hypothesis to be tested in the present studies is accordingly:

H3: Impersonal claims of certainty (e.g., it is certain) are trusted more than comparable personal claims (e.g., I am certain) and perceived to be less judgment-based.

1.4. The current studies

In the present studies participants were asked to judge statements that either contained a certainty qualifier ("I am certain" / "it is certain") or just a plain (unqualified) assertion that did not explicitly describe the strength of a belief.

Studies 1-3 were designed to explore differences in credibility. The two kinds of statements were compared directly in a within-subjects design (which speaker is more confident / which statement is more accurate / more likely to be right), or by assessing their probability or correctness on a numeric rating scale. The outcomes to be examined were binary (categorical) events and estimates of continuous quantities (amounts). For quantitative estimates we also investigated differences in precision by asking how the reported numbers should be interpreted: as an exact or approximate point prediction, or more vaguely as single boundaries of an uncertain interval (at least, at most).

Direct comparisons of certainty and plain, declarative statements might draw participants' attention to the presence and absence of the word *certain*, making this the main cue for credibility depending on their interpretation of the meaning of this term. In Studies 4-9 we used between-subjects designs where certain vs. plain statements were allocated to different participants, making the comparisons less transparent. Study 4 examined the credibility issue by asking participants about their willingness to accept the statements without further checking, and which ones needed to be examined more closely. Studies 5 and 6 zoomed in on the precision issue and investigated the judged correctness of under- and over-estimates. If certainty-statements are conceived to be *at least*-estimates they would be considered accurate even when they undershot the actual amounts. In Study 5a we asked participants about the correctness of under-estimates when phrased as certainty-statements, compared to plain, factual estimates. In Study 6a we asked, as a control, about the correctness of over-estimates. In Study 6b and 7b, we tackled the precision question from a different angle. We tested the meaning of "certain" (vs. plain) statements in the context of a classical risky choice problem, framed positively (as lives saved) or negatively (as lives lost). Two final studies (Study 8 and 9) examined claims of certainty and their basis upon evidence for predictions attributed to experts (climate scientists).

The role of the source of certainty was explored in several studies by including statements with a personal or impersonal (dummy) pronoun, expressing internal versus external certainty, respectively. We also expected that plain statements are perceived as based more strongly on objective evidence, whereas certainty-statements are perceived as judgment-based. This prediction was tested by judging speakers in brief dialogues in Study 5b. Study 8 compared plain statements with claims of personal certainty, and Study 9 included both kinds of certainty (personal and impersonal) within the same design.

2. Study 1

This was a pilot study based upon a single vignette about comparing estimates of driving distances. One speaker produced an unqualified numerical estimate of the distance between two towns, while the other speaker qualified his estimate as *certain* or *not certain*, preceded by a personal pronoun ("I am [not] certain"). Participants were asked to indicate which speaker appeared to be more confident, and also whether the numerical estimates should be read as point estimates or interval boundaries. We hypothesized that speakers who gave plain estimates could be viewed as equally or perhaps more confident than those who said that they were "certain". As for the interpretation of numeric estimates, that is, whether they should be considered single bound, approximate or exact numbers, we had no definite expectations. Some studies of framing have suggested that a substantial proportion of participants in risky-choice framing studies, will read numeric estimates, for instance, *200 people will be saved,* as indicating that "at least 200" will be saved, unless they are explicitly told how many people are *not* saved, and that the numbers are meant to be given a literal "exact" interpretation (Fisher & Mandel, 2021; Mandel, 2014).

2.1. Method

2.1.2. Participants

Overall, 320 participants were recruited via Prolific (3 participants with incomplete protocols or response time below 1.5 minutes were discarded), 161 women and 159 men, mean age 40.35 years (*SD* = 14.47); 92.8% identified themselves as native English speakers, and 58.7% held a bachelor's degree or higher. They were paid £0.80 for a 5 min survey (median completion time) that included the present vignette followed by a set of other, unrelated questions (£9.6/hr).

2.1.3 Materials and procedure

Participants received a short questionnaire intended to reveal their views on estimates that are claimed to be "certain". The introduction told participants: "This study is part of a project where we investigate the meaning of statements where an estimate is qualified by words like 'certain', 'likely' or 'uncertain" compared with factual statement where such words are not included. For instance, are the two doctors' statements below meaning the same?

"The child will recover in two days"

"I am certain that the child will recover in two days"

This practice item was intended to alert participants to the choice a speaker has when making a prediction, without prejudging their interpretations. Altogether 163 participants (51.3%) answered “Yes, they mean the same”, while 155 (48.7%) answered “No, they do not mean the same”.

They were then asked to imagine two local people in a pub in Teeford being asked about the driving distances from Teeford to two cities in the area. Alfred says: “The distance from Teeford to Upchester is 50 miles”. Bruno says: “I am not certain that the distance from Teeford to Downcastle is 50 miles” (Condition 1). In Condition 2 Bruno said “I am certain”. Participants were asked which speaker, Alfred or Bruno, that appeared to be more confident. Half of them were in addition asked what the two speakers intended to convey with their estimates, Did Alfred and Bruno mean *at most* 50 miles, *at least* 50 miles, *around* 50 miles, or *exactly* 50 miles? The other half were told that the distances to these cities were actually 40 [60] miles and asked to assess the correctness of Alfred and Bruno’s estimates on a 7-point scale from 1: Completely wrong, to 7: Completely right.

2.3. Results

2.3.1. Confidence

The participants held divergent views about the speakers’ confidence, as shown in Table 1. The speaker with an unqualified factual assertion (Alfred) appeared to a majority to be more confident than the speaker with a qualified assertion (Bruno), especially in Condition 1, where Bruno was “not certain” about his estimate. When Bruno claimed that he was *certain*, the preferences were more equal, yet a majority of 58.0% judged the speaker of the unqualified estimate to be more confident than the speaker who was certain (p = .055 by a two-tailed binomial test).

A post hoc test revealed that participants who said that the certain and plain statements in the practice trial meant the same, did not find the plain statement of Alfred more confident than the certain statement of Bruno (46.2%), but those who said that the practice statements did not mean the same regarded the certainty statement of Alfred as expressing more confidence (75.0%), χ^2 (1) = 12.87, p < .001.[1]

2.3.2. Estimate interpretations

What is meant by a distance estimate of “50 miles”? Only few participants (5-10%) thought that Alfred had a single bound (at least or at most) in mind. For Bruno, who had used a certainty qualifier, such interpretations were more common (13-26%). The speakers also

[1] This comparison was not included in our original predictions, but suggested by an anonymous referee.

differed in the around vs. exact interpretations. Qualified assertions (Bruno: I am certain) elicited for the most part around-interpretations, whereas Alfred's plain, unqualified assertions were more often believed to be exact, χ^2 (1) = 17.65, $p < .001$ (Condition 2 in Table 1). Plain assertions were deemed to mean an exact value three times more often than certain statements and interestingly, certain and uncertain statements were deemed as often to mean an "around" value, twice more often than plain statements.

Table 1. *Interpretations of plain and qualified (certain or not certain) distance estimates, Study 1 (N = 317).*

	Condition	Speaker	More confident	Interval bounds[1]		Point estimates[1]		Correctness[2]
				At least	At most	Around	Exactly	M *(SD)*
1	Plain (unqualified)	Alfred	94% (151)	5%	0%	32%	63%	2.26 (1.39)
	I am not certain	Bruno	6% (9)	10%	26%	61%	3%	5.28 (1.41)
2	Plain (unqualified)	Alfred	58% (91)	10%	0%	25%	65%	2.57 (1.56)
	I am certain	Bruno	42% (66)	3%	13%	60%	21%	2.50 (1.55)

Note: Alfred's statements in the two conditions were identical, whereas Bruno claimed to be "certain" in Condition 1 and "not certain" in Condition 2. Correctness was measured on a seven-point scale ranging from 1: completely wrong to 7: completely right.

[1]Interpretations of estimates (as interval bounds or point estimates) were done by participants in two conditions 1a (*n* = 62) and 2a (*n* = 48).

[2] Correctness ratings were performed by participants in conditions 1b (*n* = 98) and 2b (*n* = 109).

2.3.3. Accuracy

Half of the participants were told that Alfred and Bruno's estimates were inaccurate. They had either overestimated or underestimated the distance by 10 miles. Mean correctness ratings were accordingly below the midpoint of the 1-7 correctness scale. There were no significant differences between overestimates and underestimates, so accuracy ratings displayed in Table 1 (last column) were pooled. The accuracy perceptions of a factual statement and a certain statement were very similar. However, when Bruno said he was *not* certain (Condition 1), he was considered to be more right. His estimate was still inaccurate by 10 miles, but he seems to be rewarded as more correct for admitting lack of certainty.

2.4. Discussion

The results indicated that speakers who explicitly say that they are certain are not perceived to be more confident than speakers who say nothing about certainty or uncertainty. Rather, they were judged to be less confident, especially among participants who believed that certainty and plain statements meant not the same. (A perceived difference could, in principle, have implied both more and less confidence for certainty statements.) Qualified estimates were often interpreted as approximate, whereas plain estimates were assumed to be exact (and equally accurate). Thus the hypotheses about an advantage of certainty-statements (H1a and H2a) were not supported and instead we found some support for hypotheses that favor plain statements (H1b and H2b). However, we only investigated one context that featured a single, round outcome value (a 50 mile driving distances), and the certainties were all expressed in a personal, subjective format. The differences between type of statement both with respect to confidence and interpretation were not consistently significant, calling for a more comprehensive replication study.

3. Study 2

Study 1 indicated that plain, unqualified assertions were not perceived as inferior to those that were explicitly qualified as *certain*, but were believed to be expressed with an equal or higher confidence. Plain estimates were in addition perceived as somewhat more precise, and rarely given an *at least* or an *at most* interpretation. Study 2 was conducted to replicate these observations with more participants allocated to the certainty condition (and fewer to the not certain condition) and a broader selection of vignettes. Specifically, we included statements of categorical facts in addition to estimates of quantity, and compared personal vs. impersonal expressions of certainty.

The study was preregistered, reference #121169, and is available on the OSF project page, along with data and materials,

https://osf.io/2beh6/?view_only=fc7b9dc83c074655824516725bad4f2

3.1. Method

3.1.1. Participants

Participants were recruited by Prolific, and were paid £0.80 for completing the current questionnaire followed by a few unrelated questions (6 min median completion time; £8/hr). Seven participants who did not complete the relevant questions were excluded, and all the remaining cases fitted the minimum completion time of 2 minutes. So the responses from 323 participants were analyzed, 161 women, 157 men, 2 non-binary; mean age 40.6 years (*SD* = 13.71); 94.7% were native English speakers, the others considered their proficiency in English to be advanced or expert; 63.2% held a bachelor degree or higher.

3.1.2. Materials

Participants in this survey received four vignettes, two describing quantitative estimates and two addressing binary outcomes, presented in random order. All vignettes compared one speaker uttering a plain, unqualified assertion, with a second speaker who qualified their statements with “I am certain” (personal) or “it is certain” (impersonal). The sentences focused on similar but independent outcomes. Participants assessed which of the two statements convey the highest probability using different questions (more confident, higher probability, more correct, less correct). The quantitative prediction vignettes also included questions about the outcome interpretation (at least/exact/approximate). The vignettes were, as in Study 1, introduced with a practice item asking participants whether two statements with or without a certainty qualifier meant the same or meant something different from each other.

Distances. This vignette was identical to the one used in Study 1, with Alfred and Bruno estimating the driving distance between two towns. Alfred uttered an unqualified statement and Bruno an internal certainty or uncertainty phrase: “I am certain that the distance from Teeford to Downcastle is 50 miles”, or “I think the distance from Teeford to Downcastle is 50 miles, but I am not certain”. Participants were asked which speaker that appeared more confident, and were subsequently asked either (a) to interpret the meaning of these statements, or (b) to rate how wrong or right the speakers were on a 1-7 scale, after being informed that the actual distances were longer (60 miles) or shorter (40 miles) than the speakers’ estimates.

Sea level rise. The second quantitative vignette described two climate scientists predicting future sea level rise around Taiwan and Hawaii. George uttered a plain, unqualified statement: “Along the coasts of Taiwan, the sea level will rise by 50 cm”. The other speaker, Jamie, expressed certainty in impersonal terms, and said: “*It is certain* that along the coasts of Hawaii, the sea level will rise by 50 cm”. Which one of these predictions appears more trustworthy? Participants were subsequently asked to interpret the estimates (at least/at most/around/exactly 50 cm), and finally to estimate the probabilities entailed by George’s and Jamie’s predictions, along a continuous probability scale from 0 to 100.

Election result. Two political scientists are asked to predict the outcomes of two upcoming political elections. Arthur says: “In Oak County, Laura W. will win”. Ben says: “In

Maple County, *it is certain* that Jane B. will win". Which of these two experts is, in your opinion, more likely to be right?

Chess players. Two chess fans discuss the games from a chess tournament that took place some years ago. Charlotte says: "In the game between Francis and Gordon, Francis won". Diana says: *"I am certain* that the game between Harold and Ian, was won by Harold". It turns out that only one of these chess fans was correct, the other was mistaken. Who do you think was wrong, Charlotte or Diana? The answer to this question was reversed coded (for comparability with the Election vignette).

3.3. Results

Participants in the distance vignette regarded Alfred, who made the plain statement, as more confident than Bruno, even though Bruno claimed to be certain about his estimate (see Table 2), replicating the results from Study 1 with almost identical percentages (58% preference for Alfred). Due to a larger *N* this preference is significant with a binomial test (p = .012). Participants who indicated the two practice sentences did not mean the same were particularly in favor of Alfred's plain statement (64.5% vs. 51%; χ^2 (1) = 4.05, p =.044).

Regarding correctness, Bruno in Condition 1 was again "rewarded" for being not certain of his estimate. He was rated as more correct than Alfred who made an unqualified prediction, $t(72) = 6.72$, $p < .001$, *Cohen's d* = 0.79.. In Condition 2 (plain vs. certain), both speakers were rated as about equally incorrect, $t(141) = 0.32$, $p = .75$, *Cohen's d* = 0.03. Overall, the over-estimates were judged as similarly accurate as the under-estimates, for the plain statement, $M_{over} = 3.06$, $SD = 1.85$, $M_{under} = 2.74$, $SD = 1.82$ and for the un/certain statement $M_{over} = 3.45$, $SD = 1.83$, $M_{under} = 3.51$, $SD = 1.87$.

Table 2. *Interpretations of plain (unqualified) vs qualified (I am [not] certain / it is certain) distance and sea level estimates in Study 2*

Vignettes			Interval bounds		Point estimates		
Distances	Speaker	More confident	At least	At most	Around	Exactly	*M* (*SD*) correctness
Condition 1							
Plain (unqualified)	Alfred	108 (99%)	0%	3%	19%	78%	2.67 (1.85)
I am not certain	Bruno	1 (1%)	5%	5%	89%	0%	4.47 (1.40)
Condition 2							
Plain (unqualified)	Alfred	124 (58%)	4%	3%	46%	47%	3.02 (1.83)
I am certain	Bruno	90 (42%)	3%	0%	39%	58%	2.98 (1.84)
Sea level		More trustworthy					Mean (*SD*) probability
Plain (unqualified)	George	111 (31%)	16%	6%	46%	33%	83.7 (17.0)
It is certain	Jamie	212 (69%)	31%	5%	24%	40%	89.7 (16.8)

Regarding the interpretations, most estimates were believed to be approximate or exact point estimates, as shown in Table 2. The plain estimates were perceived as more exact than the "not certain" estimates in Condition 1, χ^2 (3, N = 37) = 9.43, p = .051, φ = .51, whereas both unqualified and "I am certain" estimates were about equally exact in Condition 2.

In the *Sea level* scenario, in which we asked the same questions as in the distance scenario, the speaker who qualified his statement with "it is certain" appeared more trustworthy than the speaker who made an unqualified prediction, as displayed in the bottom panel of Table 2 (p < .001 with a binomial test). The speaker of the certain statement (Jamie) was also credited with having a higher probability in mind, t(322) = 4.14, p < .001, Cohen's d = -0.23.

The difference between the distance and the climate scenarios might reflect a change from personal to impersonal certainty, and perhaps also the fact that the sea level scientists, unlike Alfred and Bruno in the distance scenario, were said to be experts in their field. Yet another reason for the certainty statement in the sea level scenario to convey a greater probability (than the factual statement) might be that the number of "at least" interpretations were considerably higher in that vignette relative to the distance vignette. This might be due to the content, which described a prediction of a gradually increasing quantity (sea level) as opposed to the geographical distances, which are assumed to be fixed values.

Table 3. *Judged correctness of plain vs qualified assertions of binary outcomes, in Study 2 and 3*

	Speaker	Type of assertion	More likely to be right*	Binomial test of differences
Study 2				
Chess winner	Charlotte	Plain	57%	*p* = .005
	Diana	I am certain	43%	
Election winner	Arthur	Plain	30%	*p* < .001
	Ben	It is certain	70%	
Study 3				
Chess winner	Charlotte	Plain	51%	Ns
	Diana	It is certain	50%	
Election winner	Arthur	Plain	58%	*p* < .001
	Ben	I am certain	42%	

*In the chess vignette, the original question: "who do you think is wrong", was reversed in coding

The two binary outcomes scenarios, *chess* and *elections*, also differed with respect to personal vs. impersonal certainty (see Study 2 in Table 3). Diana, who declared "I am certain" about the winner of the chess game, was more likely to be wrong than Charlotte, who did not mention certainty. In contrast Ben, who had announced that the election outcome was "certain"about the outcome o, was considered more likely to be *right* than his colleague who said nothing about certainty. Thus in the chess vignette, the plain assertion seemed more trustworthy than personal certainty, but in the election vignette, impersonal certainty was trusted more.

3.4. Discussion

Taken together, the four vignettes examined in Study 2 indicate that plain, unqualified assertions can be preferred to "I am certain"-statements, both for estimates of continuous quantities (distances) and binary facts (past winners), whereas "It is certain" statements were more trustworthy than plain ones both in expert predictions of future quantities (sea levels) and predictions of binary outcomes (future winners). However, the vignettes also differed along several other dimensions. For instance, the speakers in the "it is certain" vignettes were described as experts. Both these vignettes focused on predictions about future events, whereas the "I am certain"-statements described already settled, knowable events. To unconfound these differences, Study 3 was performed.

4. Study 3

Study 2 showed that statements that were explicitly claimed to be "certain" could be both more and less believable than plain assertions that did not mention certainty. Inspection of results revealed that the certainty statements were preferred over plain statements when certainty was introduced in an impersonal clause, whereas plain, unqualified assertions were preferred to certainty embedded in a personal phrase. Study 3 was conducted to investigate whether this pattern of responses would change when for each scenario personal and impersonal phrases were swapped (e.g., the distance scenario used personal certainty in Study 2, and was amended to use impersonal certainty in Study 3).

The study was preregistered reference #126226, available on the OSF project page https://osf.io/2beh6/?view_only=fc7b9dc83c074655824516725bad4f20.

4.1. Method

4.1.1. Participants

Participants who did not complete previous studies were recruited via Prolific and were paid £0.70 for a short survey (4 min median completion time, £11/hr). After discarding one respondent who did not pass the consent stage), 299 were retained for analysis (all meeting the 1.5 min minimum response time[2]); 147 men and 147 women, 5 other; mean age = 38.6 years (*SD* = 12.66); 91.0% identified themselves as native English speakers, the others described their fluency in English as advanced or expert, and 61% held the equivalent of a bachelor degree or higher, making the samples in Study 2 and 3 highly comparable while having different respondents.

4.1.2. Materials

This study presented the same four vignettes as in Study 2, but with the impersonal uncertainty (it is certain) in the distance and election vignettes replaced by personal certainty (I am certain), and vice versa in the sea level and chess vignettes. In addition, an extra quantitative vignette was introduced, describing the capacity of two different airplanes to travel 1,500 miles without refueling. This vignette focused on a more elastic measure than the distance between two towns. In the distance between towns, the distance stated is correct or not – whereas in the distance a plane can cover, under-estimates might be considered correct too, since a plane that can cover 200 km has covered 100 km to some point. This kind of scalar quantity might facilitate an *at least*-interpretation of the predicted values.

4.2. Results

[2] Note that the variation in minimum response time were based on estimates of an appropriate study completion time, which included other tasks varying in lengths across studies.

Results from the three scenarios describing quantitative estimates are summarized in Table 4 (Distance between town, Distance of travel of a plane and Sea level rise). The plain speakers were considered more confident and more trustworthy than the certain speakers both in the distance and the sea level vignettes (binomial *p*s < .001), and about equally confident concerning the flight range capacity of airplanes. Thus, the added certainty did not improve the statements' credibility, on the contrary, for two statements out of three.

Table 4. *Interpretations of plain vs qualified assertions (I am/ it is certain) of quantitative estimates in Study 3*

Scenario	Speaker	Type of estimate	More confident	At least	At most	Around	Exactly
Distance between towns	Alfred	Plain (unqualified)	58% (172)	2%	1%	70%	57%
	Bruno	It is certain	43% (127)	10%	3%	27%	60%
Plane capacity	Elena	Plain (unqualified)	51% (151)	16%	6%	36%	42%
	Sophia	It is certain	50% (148)	32%	6%	29%	32%
			More trustworthy				
Sea level	George	Plain (unqualified)	59% (175)	13%	2%	41%	42%
	Jamie	I am certain	42% (124)	28%	6%	32%	34%

The change from internal to external certainty in the distance between towns vignette did not change its trustworthiness, with preference for the plain statement being virtually the same (57.9% in Study 2 and 57.5% here). Again, those who judged that the two practice statements meant *not* the same held Alfred to be still more confident (67.1%), $\chi^2(1, N = 299) = 15.25$, $p < .001$. However, the change from external to internal certainty in the sea level vignette reduced its perceived trustworthiness, compared to the results of this vignette in Study 2, 69% vs. 42%, $\chi^2(1, N = 622) = 36.50$, $p < .0001$; Cramer's V = 0.242.

Plain estimates were in all scenarios interpreted primarily as point estimates (around or exactly the proposed number), whereas certainty-statements were more often given *at least*-interpretations, suggesting that being certain implies, perhaps paradoxically, estimates that are less precise, but have the potential to cover increasingly large outcomes.

The chess and the election binary outcome vignettes showed a slight preference for plain statements as being more likely to be correct (51% and 59% respectively, as shown in the lower panel of Table 3). Participants' preferences in those vignettes appeared to be affected by the change in type of certainty. When predictions of election outcome were changed from *it is certain* (in Study 2), to *I am certain* in the present study, the prediction became less credible /70% vs. 42%). A comparison of frequencies reported in the upper and lower panel of Table 3 yielded a highly significant effect for the election vignette, χ^2 (1, N = 622) = 46.71, p < .0001; Cramer's V = 0.27). Consistently, in the chess vignette, which was changed from personal to impersonal certainty, the certainty statement became more likely to be correct, although this change was not statistically significant; 43% vs. 50%, χ^2 (1, N = 622) = 2.87, p = .09, Cramer's V= 0.07.

4.3. Discussion

Estimates of numerical values, predictions and assertions about past outcomes did not become more believable when speakers claimed that they were certain. This may be related to another finding from Study 1 and 2, namely that certainty statements are more often interpreted as referring to a vague, potentially larger outcome, than plain, unqualified statements. The outcome value described as being certain were more often taken as representing the lowest bound of an open range (e.g., at least 200km). Results of Study 3 confirmed the conjecture that personal certainty is trusted less than certainty formulated in impersonal terms, in line with Hypothesis 3. Speakers who announce "I am certain" are trusted less, perhaps because the personal pronoun indicates that other speakers might be of a different opinion, and disagreement among experts appears to be a particularly pernicious source of doubt in science (Gustafson & Rice, 2020). "It is certain" suggests, in contrast, objectivity and consensus. Yet in most instances, we found that it is better not to mention certainty at all.

5. Study 4

In the three preceding studies, participants were asked to compare certainty-statements and plain statements presented jointly, in a within-Subjects design. It appeared that plain statements were generally preferred but not in all cases. For instance, in the Sea level vignette in Study 2, an expert who said that it was certain was considered more trustworthy than one who did not mention certainty, and in the Plane capacity vignette in Study 3 certainty and plain speakers were judged to be equally confident. Participants in all these studies judged statements with certainty included or omitted presented side by side. Joint presentations will often serve to highlight essential differences (Birnbaum, 1999; Hsee, 1998), but on the other hand they may lead to judgments based on less important, but salient distinctions (Hsee &

Zhang, 2004). For instance, some participants may have inferred that absence of the term *certain* in the plain assertions could imply that these claims were *not* certain**,** which obviously made them less confident and less trustworthy. The present study followed instead a between-subjects design, where participants in separate conditions rated their acceptance of unqualified and certainty statements as occurring in a normal conversation. For instance, if a friend tells you "Jack owns two cars", would it be reasonable to ask for additional evidence, or would you be happy to accept this piece of information as a fact that does not need to be substantiated further? Our hypothesis was that plain statements are less in need of being checked than corresponding certainty-statements. If Jack "surely owns two cars", his ownership seems a bit contentious.

5.1. Method

5.1.1. Participants

Altogether 301 participants were recruited via Prolific as in the previous studies, 148 men and 148 women, 5 other; mean age = 38.6 years (SD = 12.4); 92.4% identified themselves as native English speakers, the others described their fluency in English as advanced or expert, and 68.8% held the equivalent of a bachelor degree or higher.

5.1.2. Materials and procedure

Participants received three blocks of questions. Block 1 contained eight statements which were introduced as follows.

> "Imagine that, in the course of a conversation, one speaker issues the following statement: "The Eiffel tower is more than twice as high as Big Ben".
>
> Would you normally accept the statement at face value (and be willing to retell it as a fact)? Or would you be tempted to ask for supporting evidence, for instance by asking "how do you know"?
>
> You will now read a series of 8 statements that could occur in a normal conversation. For each statement, please evaluate whether they need to be checked or if they don't need to be checked and can be taken as a fact."

We developed a list of 8 statements focusing on quantitative values (e.g., how many mistakes made in a test). There were two versions of each statement: One plain assertion (e.g., Henry made 4 errors in his test) or a certain one (e.g., It is certain Henry made 4 errors in his test). Certainty was expressed in several ways across statements, with personal pronouns (I am certain), impersonal dummy pronouns (I am certain) or other explicit assurances, such as "for sure" and "definitely". For complete list of the eight statements, see Table 5. Participants in Condition 1 received the plain versions of statements 1-4, and the certainty versions of statements 5-8. Participants in Condition 2 received the opposite versions (presented in a randomized order to each participant). Each statement was judged on a four point scale for acceptability (1: Needs checking, 2: Might need checking, 3: Might not need checking, 4: Does not need checking).

Block 2 was based on the Chess and Election vignettes from Study 2 and 3. The vignettes focused on a (past or future) categorical outcome as before but were improved in two ways. First, together with a plain statement, the study included either a personal or an impersonal certainty statement (between-subjects) to enable a more direct comparison. Second, we homogenized the two vignettes by making them both focus on either success or error.

So for example, in the election vignette, participants could read the following with experimental variations between square brackets:

Arthur says: "In Oak County, Laura W. will win".

Ben says: "In Maple County, *it is certain* [*I am certain*] that Jane B. will win".

Which of these two experts is, in your opinion, more likely to be right [one of the two is mistaken. Who is more likely to be wrong]?

Block 3 explored the suitability of unqualified vs. certainty statements for correcting inaccurate estimates. In two conditions, participants were asked to correct a wrong plain estimate (Condition 1) and a wrong certainty statement (Condition 2). Would they prefer to correct these statements with precise or with approximate numbers? One statement was about the number of Labour representatives in the Parliament, a second about the length of the river Thames, the third concerned the date for the Russian invasion in Ukraine. We assumed that inaccurate plain estimates would be corrected with exact values whereas inaccurate certainty estimates would more likely be corrected with approximate values.

5.2. Results

All plain, unqualified statements in Block 1 were judged to be more acceptable than the corresponding certainty statements.[3] For instance, few would question "Jack owns two cars", whereas "Jack surely owns two cars" should be checked more closely before accepted.

[3] Due to a clerical error, statement 8 (Lola spent five hours on the job) contained the term "certainly" in both conditions. This statement was therefore removed from the subsequent analyses.

In other words, claims of certainty made statements more questionable than statements without such claims. As shown in Table 5, the difference was highly significant for six of the seven statements tested, with medium to large effect sizes. Altogether, plain statements were mostly rated 3 or 4 (64.7% of all ratings) - i.e., a majority thought they need not be checked, whereas the corresponding certainty statements received only 42.7% such ratings, i.e., a majority thought they should be checked.

Table 5. *Mean ratings (1-4) of acceptability for statements without and with information about certainty in Study 4* (1: Needs checking, 4: Does not need checking)

Unqualified statements	Acceptability	Certainty-statements	Acceptability	Tests of difference (df =299)	*Cohen's d*
Anna has three kids	3.46	Anna certainly has three kids	3.07	t = 3.66, p < .001	0.42
Jack owns two cars	3.24	Jack surely owns two cars	2.29	t = 8.82, p < .001	1.02
Bridget paid £300 for her bike	3.05	I am certain that Bridget paid £300 for her bike	2.37	t = 6.41, p < .001	0.74
Kenneth's income will be £5,000 next month	2.36	Kenneth's income is guaranteed to be £5,000 next month	2.21	t = 1.22, p = .144	0.14
Clara is 40 years old	3.18	Clara is definitely 40 years old	2.70	t = 4.54, p < .001	0.52
The temperature was 20^{o}C yesterday at noon	2.86	I am certain that the temperature was 20^{o}C yesterday at noon	2.23	t = 5.90, p < .001	0.68
Henry made four errors in his test	2.63	It is certain that Henry made four errors in his test	2.15	t = 4.13, p < .001	0.48
Lola spent five hours on that job[1]	missing data	Lola certainly spent five hours on that job	2.36	not applicable	n.a.

Note: Participants in Condition 1 (*n* = 151) rated the plain versions of statements 1-4 and the C-versions of statements 5-8. Participants in Condition 2 (*n* = 150) rated the complementary set of statements.

[1]Participants in this condition erroneously received the certainty statement (no difference between the two conditions).

Block 2 replicated the two binary outcome vignettes (chess and election) from Studies 2 and 3 (see Table 6). Charlotte's unqualified statement was considered more likely to be right than Diana's certainty-statement. This difference was reduced when Diana used an impersonal phrase (it is certain)both in the positive frame, χ^2 (1) = 6.12, p = .013, and in the negative frame, χ^2 (1) = 6.05, p = .014.

For the election vignette, we found a similar trend, but only in the negative frame, χ^2 (1) = 6.46, p = .011. In the positive frame of the election vignette the two speakers were perceived as equally likely to be right.

Table 6. *Judged correctness of unqualified vs. (I am / it is) certainty assertions of binary outcomes, in Study 4, Block 2*

Scenario	Speaker	Type of assertion	Plain vs. certain more likely to be correct	Binomial test of differences plain vs. certain
Positive frame – who was right?				
Chess	Charlotte vs. Diana	Plain vs. I am certain	102 (67.5%) vs. 49	$p < .001$
		Plain vs. It is certain	81 (53.6%) vs. 70	*ns*
Election	Arthur vs. Ben	Plain vs. I am certain	76 (51.0%) vs. 73	*ns*
		Plain vs. It is certain	73 (48.3%) vs. 78	*ns*
Negative frame (R)* Who was wrong?				
Chess	Charlotte vs. Diana	Plain vs. I am certain	98 (65.3%) vs. 52	$p < .001$
		Plain vs. It is certain	77 (51.3%) vs. 73	*ns*
Election	Arthur vs. Ben	Plain vs. I am certain	86 (57.0%) vs. 65	$p = .052$

Plain vs. It is certain	63 (42.3%) vs. 86	*p* = .036

* In the chess vignette, the original question: "who do you think is wrong", was reverse coded to make the results comparable across question frame.

In Block 3 we examined whether a wrong estimate should be corrected by a precise or more approximate number depending on whether it was presented as certain. The number of Labour MPs and the length of the river Thames were preferably corrected with a known, precise value. This applied especially to plain estimates (66.7% in the parliament vignette and 77.2% in the Thames vignette), but also to certainty estimates (60.3% and 69.5%, respectively). All these preferences are significantly different from 50% with binomial tests (p < .001), but not significantly different from each other. The Russian invasion in Ukraine led to a more even distribution of answers, with no significant preference for corrections in terms of date (precise), or month (approximate).

5.3. Discussion

Results from the first block indicated that plain statements in the context of an ordinary conversation are generally not questioned, but accepted as factual pieces of information, but when the same statements were qualified as certain, it became reasonable to examine them more closely, for instance by asking "how do you know". We did not distinguish between personal and impersonal certainty in this study, and most statements on the list were of an impersonal kind.

Results from Block 2 (direct comparison of speakers using plain and I vs. it certainty-statements) replicated findings from Study 3 with respect to impersonal vs personal certainty. Statements of impersonal certainty were taken as more likely to be correct than personal ones. The two vignettes differed in preference for plain statements (more marked preference in the chess scenario), perhaps because of a difference in time frame (past vs. future focus), whereas the outcome correctness frame (likelihood of being right vs. who is wrong) appeared to be of minor importance. One may speculate that certainty about events that have happened can be attributed to people's fallible memory, and could be more of an epistemic kind, whereas certainty about events that are going to happen in the future is more strongly based on external evidence, like opinion polls.

Results from Block 3 indicated that people think that wrong numeric estimates should be corrected by point estimates rather than by open ended interval estimates, even if the latter has an implicit directionality (*nearly* and *more than* indicate that we are talking about large numbers; Teigen, 2022).

6. Study 5

The sea level and plane capacity vignettes in Study 2 and 3 suggested that certainty statements allowed for more at least-interpretations than just plain, unqualified estimates that were more taken as approximate or exact values. Study 5 was designed to test this notion in an alternative way, by using the list of eight statements estimating quantities used in Study 4 along with information about the correct values. For instance, when a speaker says: “Henry made four errors on the test”, and it turns out that Henry actually made five errors, how correct or incorrect is the speaker’s estimate? From an *at least*-interpretation of numbers, the estimate may be regarded as correct, but from an *exact* interpretation, the estimate is obviously wrong. We expected that if certainty-statements permit more *at least*-interpretations than corresponding unqualified statements, they would also be judged as more correct in cases where the true values are higher. So if Henry actually failed five items on the test, it is incorrect to say that he made four errors, but more admissible to say that it is *certain* he made four errors, in line with Hypothesis 2b (i.e., statements qualified by “certain” convey a *less* precise and reliable estimate than plain statements).

Numeric statements that allow for *at least*-interpretations suggest the lower bounds of intervals whose upper limits are not specified. They are accordingly more vague and approximate than point estimates, and may be based on subjective judgmental processes (memories and intuitions) rather than on objective sources of knowledge (measurements, instrument readings). Study 5 included three brief dialogues where one speaker produced either a plain estimate or a certainty-statement, the question being whether these estimates were derived from objective observations or if they reflected the speaker’s subjective judgments. For instance, if a speaker estimates the time to be “three o’clock”, or “certainly three o’clock”, does this imply that she has checked the time, or is she guessing? We hypothesize that an unqualified point estimate would require access to a watch. By mentioning degree of certainty, the speaker admits that the estimate is based on personal judgments and is approximate.

The study was preregistered reference #133073, available on the OSF project page along with data and materials https://osf.io/2beh6/?view_only=fc7b9dc83c074655824516725bad4f20.

6.1. Method

6.1.1. Participants

Participants were recruited via Prolific as in the previous studies (and those who participated in Study 1-3 were not allowed to take part). They were paid £0.90 for a 5 min study (median completion time; £10.6/hr). After discarding data from one respondent who did not consent and did not answer any questions, 301 were retained for analysis (response time > 1.5 min); 147 men and 152 women, 2 other; mean age = 41.0 years (*SD* = 13.1); 92.4% identified themselves as native English speakers, the others described their fluency in English as advanced or expert, and 67.1% held the equivalent of a bachelor degree or higher. Participants were randomly allocated to one of two conditions for the first part of the questionnaire (Block 1: the list of statements), and to one of three conditions for the second part (Block 2: dialogue scenarios).

6.1.2. Materials and procedure

Block 1: How correct are the (over)estimates? Participant received eight statements, identical to those used in Study 4 (block 1), with or without a certainty term, for instance, “Ann has [certainly] three kids” and “Clara is [definitely] 40 years old”. They were also informed about the true amount, which in all cases were higher than the stated value (Ann had actually four kids; Clara was in fact 45 years old) and were asked to evaluate the correctness of the original statements on a four-point scale (1: Not correct; 2: Somewhat incorrect; 3: Somewhat correct and 4: Correct). For a complete list of statements, see Table 7. Participants in Condition 1 received the plain versions of statements 1-4, and the certainty versions of statements 5-8. Participants in Condition 2 received the opposite versions (presented in a randomized order to each participant)

Block 2: Which estimates are based on evidence? This block included three scenarios that featured an interlocutor who answered a question with either a certainty statement (internal or external), or with a plain numeric estimate (random between-subjects allocation).

Scenario 1. Jack and Jill are awakened by outside noises in the middle of the night. Jack asks: “What time is it?” Jill answers one of the following three statements (randomly allocated): (i) “It is three o’clock”, (ii) “It is certainly three o’clock”, or (iii) “I am certain it is three o’clock”. Question 1. Did Jill check he time or did she guess? She checked / She guessed. Question 2. Jack grabs his watch to check the time. He grunts: “Oh well, close enough, the time is actually 2:50 am/ 2:55 am / 3 am / 3:05 am / 3:10 am” (select one).

Scenario 2. Dina and Dave is on summer holiday and would like to go for a swim. Dina asks about the water temperature and Dave answers one of the following three statements (randomly allocated): (i) "It is 20 degrees", (ii) "It is 20 degrees for sure", or (iii) "I am sure it is 20 degrees". Question 1. Did Dave check the water temperature or did he just guess? Question 2. Arriving at the beach, Dina reads about the water temperature on the beach notice board: "You were quite right, the water is 18 / 19 / 20 / 21 / 22 degrees" (select one).

Scenario 3. Julie and Paul have an argument about how much alcohol there is in a local drink. Julie thinks there is barely any alcohol in it. Paul says one of the following three statements (randomly allocated): (i) "It is 20% alcohol", (ii) "It is certain that it is 20% alcohol", or (iii) "I am certain that it is 20% alcohol". Question 1. Did Paul read this on the bottle label, or did he guess? Question 2. Julie buys a bottle of the local drink and examines its label. Then she says: "You were fairly correct. It says that it is 16 / 18 / 20 / 22 / 24% alcohol" (select one).

6.2. Results

6.2.1. Block 1: Incorrect statements

All eight statements underestimated the correct values (listed in the column for "True values" in Table 7. For instance, Henry was reported to have made four errors, although he had actually made five. Most plain statements were accordingly rated to be incorrect or somewhat incorrect (71.2% of all ratings). However, certainty statements were perceived as more correct, with 51.2% rated (somewhat) incorrect and 48.8% (somewhat) correct. Mean correctness ratings for each item are displayed in Table 7, showing that certainty-statements were viewed as significantly more correct than plain statements in six out of eight pairs, whereas two pairs did not differ significantly. These results support that plain, unqualified statements are generally given exact interpretations, whereas estimates qualified with "certain" and certainty equivalents (certainly, surely, definitely, guaranteed) are more often given *at least*-readings, and hence perceived to be correct even when they understate the true amounts.

Table 7. *Mean ratings (1-4) of correctness for under-estimate statements with and without information about certainty, indicating that people endorse more a vague interpretation of certainty-statements (Study 5, Block 1)*

Plain statements	True values	Correctness of plain statements	Certainty-statements	Correctness of certainty-statements	Tests of difference (*df*=299)	*Cohen's d*

Anna has three kids	4 kids	1.71	Anna certainly has three kids	2.35	*t* = 5.48, *p* < .001*	-0.63
Jack owns two cars	3 cars	2.05	Jack surely owns two cars	2.53	*t* = 4.25, *p* < .001	-0.89
Bridget paid £300 for her bike	£350	1.68	I am certain that Bridget paid £300 for her bike	1.83	*t* = 1.48, *p* = .141	-0.17
Kenneth's income will be £5,000 next month	£5,500	2.20	Kenneth's income is guaranteed to be £5,000 next month	2.97	*t* = 6.97, *p* < .001	-0.80
Clara is 40 years old	45 years	1.35	Clara is definitely 40 years old	1.79	*t* = 4.52, *p* < .001*	- 0.52
The temperature will be 20°C tomorrow at noon	22°C	2.23	I am certain that the temperature will be 20°C tomorrow at noon	2.13	*t* = 0.97, *p* = .335	0.11
Henry made four errors in his test	5 errors	1.74	It is certain that Henry made four errors in his test	2.34	*t* = 5.17, *p* < .001*	-0.60
Lola spent five hours on that job	6 hours	1.85	Lola certainly spent five hours on that job	2.61	*t* = 7.22, *p* < .001	-0.83

Note: Participants in Condition 1 (n = 152) rated the plain versions of statements 1-4 and the certainty versions of statements 5-8. Participants in Condition 2 (*n* = 149) rated the complementary set of statements.

* Equal variance not assumed.

6.2.2. Block 2. Dialogues

In all three dialogues, speakers who said they were certain were generally believed to be guessing, rather than having checked the actual values. Speakers who uttered unqualified, plain estimates appeared to be better informed, as indicated by the percentages of Checked vs Guessing responses reported in Table 8. Speakers expressing impersonal certainty ("it is certain") fared somewhat better than those who described their personal certainty ("I am certain"), especially in Scenario 3.

The correct values suggested by the second speakers in the dialogues were generally slightly lower than the estimates stated by the first speakers: The time of night was on average 7.2 minutes before 3am, and the alcohol percentages in the local drink was 19.1% rather than the estimated 20%. For water temperatures 20 degrees was judged to be an accurate estimate. The actual values of the plain and qualified statements were similar (no significant differences).

Table 8. *Percentages of participants in three conditions who believed that speakers' estimates were based on objective information (they checked) or unaided subjective judgments (they guessed), in three dialogue scenarios (Study 5, Block 2).*

Scenario	Plain statements		It is certain (for sure)		I am certain (I am sure)		Tests of significance
	Checked	Guessed	Checked	Guessed	Checked	Guessed	
Time of night	70.0	30.0	38.0	62.0	27.7	72.3	$\chi^2(2)= 39.38$ $p < .001$
Water temperature	33.3	66.7	10.9	89.1	6.9	93.1	$\chi^2(2)= 28.72$ $p < .001$
Alcohol in drink	53.5	46.5	58.2	41.8	20.6	79.4	$\chi^2(2)= 34.36$ $p < .001$

6.3. Discussion

For all eight statements in Block 1, true values were selected to be numerically higher than the speakers' estimates. All the statements were accordingly, if taken literally, understatements. Nevertheless, and consistent with an *at least*-interpretation, certainty-statements were more often considered right than wrong, whereas unqualified statements were judged to be more wrong than right. In other words, certainty-statements, both of a personal and an impersonal kind, are readily given an *at least*-reading, whereas plain statements are mostly viewed as exact or approximate point estimates. In contrast, the dialogues in Block 2 indicated that a speaker's estimates were assumed to be slightly above rather than below the professed true values (both for certainty and plain statements). This unpredicted finding may be due to our instructions in the dialogue partners' comments about the speakers' original estimates as "close enough", or "fairly correct", which conveyed the notion that the values were not spot on, but "almost there", just below the correct value. *Almost* and its equivalents typically denote an estimate that is approaching the stated value from below (Ferson et al., 2015; Kahneman & Varey, 1990; McKenzie & Nelson, 2003).

We did not in this study collect data on the correctness of overestimates, (e.g., Henry is reported to have made five errors but actually has made only four). We predicted that underestimated certainty statements would be rated as less correct than unqualified statements (see Study 6).

The dialogues in Block 2 revealed that paradoxically, claims of certainty can indicate that the speakers are guessing instead of basing their estimates on objective evidence. This is consistent with results from the preceding studies indicating that speakers who express certainty are sometimes less certain than those who do not mention certainty. This was especially the case for personal (I am certain) certainty statements. It also follows from an *at least*-reading of certainty implying an interval (as opposed to point estimates), and more generally a "gist" interpretation suggesting that certain means "a lot" (H2a). Alternatively, an explicit claim of certainty may pragmatically implicate the existence of alternative but rejected estimates – which might be under- or over-estimates (as in dialogues where the conversation partners question the estimates or are of a different opinion).

The link between plain statements and evidence works both ways. Undergraduate students in a Judgment and Decision-Making class were given the dialogue scenarios and asked to play the role of Jill. If they had a watch to look at, 20 out of 25 (80%) would use the plain option: "It is three o'clock". If they had no watch to look at, 19 out of 25% (76%) would answer "I am certain it is three o'clock" rather than the other two alternatives.

7. Study 6

The list of statements rated for correctness in Block 1 of Study 5 described estimates that were consistently *lower* than the true values. The rated correctness of certainty-statements was compatible with an at least-interpretation of such statements. However, it could alternatively mean that certainty statements were perceived as approximate rather than exact, and might hence be judged as compatible with both higher and lower true values. To control for this possibility, we presented in Study 6 the same list of statements, but adjusted the true values downwards. For instance, Ann is said to have [certainly] three kids, when she in fact has only two. We predicted in this case that certainty statements would be rated as less correct than in the prior study, and perhaps more wrong than plain, factual statements.

This study was also designed to examine the effects of *certain* upon at least-interpretations in risky framing problems. In the seminal Asian Disease scenario, Tversky and Kahneman (1981) showed that people preferred a program expected to save 200 people out of

600 whose lives were at stake, rather than a program that had a 1/3 chance to save them all. Yet, most participants preferred the "risky" option when the programs were framed in terms of lives lost rather than lives saved. Losing 400 out of 600 lives seemed worse than a 1/3 probability of no lost lives. These framing results have been replicated and discussed for more than forty years. One frequently voiced criticism is that the options are insufficiently described and hence not formally equivalent, so the shift needs not imply that people's preference are irrational (Frisch, 1993; Kühberger & Tanner, 2010, Okder, 2012). Geurts (2013) and Mandel (2001, 2014) claimed that saving 200 people is ambiguous, as it could mean 200 *or more.* Correspondingly, 400 lives lost could be read as 400 *or more.* Such *at least*-readings of numbers has been claimed by linguists to be fairly common (Breheny, 2007; Spector, 2013). This obviates the equivalence of the two programs and makes the positively framed option more appealing than the negatively framed version also in an objective sense. Saving at least 200 lives is clearly better than losing at least 400 of the 600 lives at stake.

The "riskless" program, framed positively as 200 lives saved and negatively as 400 lives lost, has often been described as the *certain* option, but was originally presented to participants as plain, unqualified predictions, no mention being made of certainty (Tversky & Kahneman,1981). Mandel added "it is certain" to his version of the problem, explaining that this was done to make "the assumed certainty of stated options explicit" (2014, p. 1189). Similarly, Chick et al. (2016) described the riskless program as saving 200 people "for sure". This added mention of certainty would, in our view, not serve to make the stated option more certain and precise (as suggested by Mandel), but might instead have increased the availability of lower bound interpretations. We do not know any risky framing studies comparing people's choices among unqualified statements of number of lives with statements where the quantities are explicitly qualified as "certain" or "sure". If certainty-statements are more likely to imply *at least*-interpretations, we should expect such statements to facilitate rather than counteract framing effects.

7.1 Method

7.1.1. Participants

Participants were recruited by Prolific as in the previous studies (and participants from previous studies were not eligible to take part). They were paid £0.75 for a 5 min study (median completion time). All of the 298 study completions were deemed valid, and were retained for analysis (completion time > 1.5 min), 146 men and 145 women, 7 other or preferred not to say; mean age = 40.6 years (*SD* = 14.1); 92.6% identified themselves as native English speakers, the others described their fluency in English as advanced or expert, and 65.4% held the equivalent of a bachelor degree or higher. Participants were randomly allocated to two conditions for the first part of the questionnaire (the list of statements), and to four conditions for the second part (risky choice scenario).

7.1.2. Materials and procedure

The study consisted of two blocks of question. In Block 1 participants received the list of eight statements used in Study 5, but with the stated values higher than the correct ones, so that all statements to be judged were over-estimates rather than under-estimates. For instance, a speaker says that Ann has three kids when she actually has only two. The accuracy of each statement was judged as before on a scale from 1: Incorrect to 4: Correct.

In Block 2 participants received a risky choice scenario where they were asked to choose between a "certain" and a risky plan in a situation where 600 lives were at stake. The original Asian Disease scenario (which seemed inappropriate, especially after the COVID-19 pandemic), was replaced by a man-made disaster version developed by Mandel (2001, 2014). Participants were asked about their program preference and their range estimates of the lives that would be saved or lost.

In a war-torn region, the lives of 600 stranded people are at stake. Two response plans with the following outcomes have been proposed.

If Plan A is adopted, [it is certain that] 200 people will be saved (400 people will die) If Plan B is adopted, there is a one-third probability that all 600 will be saved (nobody will die) and a two-thirds probability that nobody will be saved (all will die).

- Which plan would you favor, Plan A or Plan B?
- Of the 600 people who are at risk here, how many lives might be saved (lost) by the plan you have chosen?

With this plan we can expect to save (lose) between …. and …. lives.

Participants were randomly allocated to one of four conditions in a 2 x 2 between-subjects design, with *frame* (lives saved or lost, as indicated in the parentheses) as one factor, and *certainty* (mentioned or not mentioned, as indicated in the brackets) as the second factor.

7.2. Results

7.2.1. Block 1: Incorrect statements

All eight statements listed in this block predicted quantities higher than the true values. The estimates were accordingly judged to be inaccurate (93.7% of ratings were 1 or 2 on the four-point rating scale, where 1 was "incorrect"). The ratings were similar for plain statements and certainty statements, with M = 1.35 for and M = 1.36 respectively. These ratings of over-estimates were consistently lower than the corresponding ratings for under-estimates in Study 5, displayed in Table 6. This in line with previous findings of over-statements being judged more incorrect than under-statements (Teigen & Filkuková, 2011; Teigen & Nikolaisen, 2009). More important, the statements were not in this study rated more correct by describing them as certain, implying that that the correctness of inaccurate estimates is only improved when they are *lower* than the true values, and may be given *at least*-interpretations.

7.2.2. Block 2. Risky choice scenario

Replicating classical framing studies (Steiger & Kühberger, 2018; Tversky & Kahneman, 1981), two thirds of the participants (66.7%) preferred the “safe” Plan A in the positive framing condition, against only one third (35.5%) in the negative condition, $\chi^2(1) = 34.89$, $p < .001$, $\varphi = 0.34$. The certainty quantifier slightly magnified the preference for Plan A vs. B in the positive condition, but this was not statistically significant, $\chi^2(1) = 0.87$, $p = .351$, $\varphi = -0.08$. The certainty statement also slightly increased the preference for the Plan A in the negative condition (over the uncertain Plan B) but the difference was not statistically significant, $\chi^2(1) = 0.66$, $p = .416$, $\varphi = -0.07$. Thus we cannot conclude that the word *certain* magnified the framing effects in this study.

Yet certainty affected the expected number of people saved among those preferring option A. In the plain condition a majority of 60.9% suggested an “exact” interpretation (they answered that “between 200 and 200” would be saved). In contrast, the distribution of people expected to be saved in the certainty condition was bimodal, with only 40.4% suggesting an exact reading of the estimates, $\chi^2(1) = 4.10$, $p = .043$, whereas 38.5% read the estimate to mean that at least 200 people would be saved (between 200 and 600 lives), against 8.7% in the plain condition , $\chi^2(1) = 11.70$, $p < .001$.

7.3. Discussion

Under-estimates in Study 6 (Block 1) were judged to be much more incorrect than the corresponding over-estimates in Study 5. This was especially the case for statements that were certain (vs. plain statements), consistently with an “at least” interpretation of the outcome they qualified.

The perceived correctness of over-estimates is in line with previous findings (e.g., Teigen & Nikolaisen, 2009), and is supported by related judgment asymmetries, for instance the phenomenon of additivity dominance (Scott & Rozin, 2017), whereby taking something away from a product seems to change its nature less than adding something. This stands in contrast to the negligible asymmetries in Studies 1-3 where underestimated and overestimated distances were judged to be about equally wrong. The reason may be sought in a subtle difference between a measurable stable feature (distance between towns) in the first three studies, and countable amounts that are recorded over time, where higher numbers entail smaller ones (number of kids, cars, errors) in Studies 4-6.

Answers from the risky framing task (Block 2) revealed that most participants who chose option A in the plain condition gave a single number (200) twice, despite being explicitly instructed to define a range ("between … and … lives saved"). Certainty-statements were perceived to be more inclusive, with 200-600 ranges mentioned equally often as zero ranges. However, since participants were only asked to estimate expected ranges for the plan they favoured, we do not know whether their plan preferences were derived from the expected number of lives to be saved/lost, or the other way around. This limitation was overcome in the following study where all participants were asked about their interpretations of Plan A before indicating which plan they favoured.

8. Study 7

Altogether, the previous studies indicated that plain, unqualified statements tended to be endorsed as more persuasive, less vague and more likely to be based on objective evidence, compared to statements that were explicitly qualified as certain. One reason for this perhaps surprising finding could be that unqualified statements, being objectively simpler, are more straightforward, easier to understand and require less mental processing capacity. The introduction of a certainty qualifier might introduce a level of abstraction and complexity to a statement otherwise concrete. Construal level theory (CLT) posits that people can construe the world as more or less psychologically distant, characterised by how concretely it can be represented, with concrete representations marking a low level of construal and abstract representations marking higher levels of construal (Trope et al., 2007; Trope & Liberman, 2010). Low-level construals are also perceived to be more probable than high-level construals (Wakslak & Trope, 2009) and people tend to judge low construal messages – i.e., more concrete messages - as being more true (Hansen & Wänke, 2010). Higher levels of construal also made imprecise outcomes more acceptable than lower level construal (Onay et al., 2013). Based on this overall evidence, we suggest that construal may play a role in the perceived greater precision and degree of certainty that people associate with plain (vs. certainty) statements. Block 1 of Study 7 was designed to explore the possibility that plain statements are perceived as more concrete reflecting a lower level of construal, which might provide an explanation for the perceived precision (vs. vagueness) of such statements.

Study 7 also included a set of questions (Block 2) related to the risky choice scenario used in Study 6, but participants were this time asked about the expected outcome of the riskless option (Plan A) *before* they made their choice about which plan they favoured. We expected to find effects of framing, especially among participants in the certainty condition who were predicted to suggest *at least*-interpretations of the number of lives (200-600) that might be saved.

The study was preregistered (AsPredicted, registration #137947) along with data and materials, https://osf.io/2beh6/?view_only=fc7b9dc83c074655824516725bad4f20.

8.1. Method

8.1.1. Participants

Participants were recruited by Prolific as in the previous studies. All of the 301 study completions were deemed valid and were retained for analysis (completion time > 1.5 min); 150 men and 146 women, 5 other or preferred not to say; mean age = 39.5 years (*SD* = 13.1); 95.3% identified themselves as native English speakers, the others described their fluency in English as advanced or expert, and 66.1% held the equivalent of a bachelor degree or higher. Participants were randomly allocated to two conditions for the first part of the questionnaire (the list of statements), and to one of 2 x 2 conditions for the second part (risky choice scenario), with positive vs negative frame and plain vs. certain statements as two between-subjects factors.

8.1.2. Materials and procedure

Block 1: Concreteness judgments. This block included a list of eight pairs of statements, identical to those used in Studies 4-6, but this time the same participants read both the plain and certain versions of each statement presented on the same page and judged which of the two was more concrete. Participants were allocated to two conditions that differed only in the presentation order of statement pairs, and answers from both were pooled in the subsequent analyses. Participants were initially informed that statements may vary in concreteness. For instance, the description of a table can be more concrete than the description of a theory.

Block 2: Risky framing .This part contained four versions of the risky choice framing problem as used in Study 5 and 6. Participants in the positive condition were asked how many would be saved by program A (0-200, exactly 200, 200-600, 0-600, or a self-defined range). In the negative condition they correspondingly had a choice between lives expected to be lost by implementing program A (0-400, exactly 400, 400-600, 0-600, or self-defined). Of these alternatives, 200-600 (in the positive frame) and 400-600 (in the negative frame) indicated *at least*-interpretations, and were singled out for special analyses.

8.2. Results

8.2.1. Block 1: Concreteness judgments

As shown in Table 9, in five of eight statement pairs, plain statements were judged to be more concrete than certainty statements. Four of these differences were statistically significantly different from 50%, by binomial tests. For instance, "Jack owns two cars" was judged to be more concrete than "Jack surely owns two cars" by 82% of the respondents. This trend was reversed in two statements where "guaranteed" and "definitively" were used as a proxy for certainty, and also in one statement expressing impersonal certainty ("it is certain").

Table 9. *Selection rate of the unqualified statement as more concrete than certainty statements (Study 7, Block 1, N = 301)*

Plain statements	Certainty-statements	% selection of plain statement as most concrete (n)	Binomial test of difference
Anna has three kids	Anna certainly has three kids	56% (170) vs. 44%	p = .028*
Jack owns two cars	Jack surely owns two cars	82% (248) vs. 18%	p < .001
Bridget paid £300 for her bike	I am certain that Bridget paid £300 for her bike	56% (170) vs. 44%	p = .028
Kenneth's income will be £5,000 next month	Kenneth's income is guaranteed to be £5,000 next month	23% (69) vs. 77%	p < .001
Clara is 40 years old	Clara is definitely 40 years old	38% (114) vs. 62%	p < .001
The temperature will be 20°C tomorrow at noon	I am certain that the temperature will be 20°C tomorrow at noon	56% (168) vs. 44%	p = .050
Henry made four errors in his test	It is certain that Henry made four errors in his test	41% (124) vs. 59%	p = .003
Lola spent five hours on that job	Lola certainly spent five hours on that job	51% (155) vs. 49%	p = .645

8.2.2. Block 2: Risky choice scenario

In the risky choice task, plan A was given more often an *at least*-interpretation (200-600 and 400-600 ranges) in the certainty condition than in the unqualified condition, as shown in Figure 3. There was a greater frequency of *at least*-interpretations of certainty-statements than of plain statements both for lives saved and lives lost; $\chi^2(1) = 16.67$, $p < .001$, $\varphi = 0.33$ (positive frame), and $\chi^2(1) = 7.48$, $p = .006$, $\varphi = 0.22$ (negative frame). Conversely, there were more *exact*-responses in the unqualified condition than in the certainty condition, $\chi^2(1) = 6.33$, $p = .012$.

In both conditions, plan A was preferred to plan B by a majority in the positive frame (69.3% and 78.7% for unqualified vs certain statements, respectively), but only by a minority in the negative frame (30.7% and 30.3%), replicating the framing effect from Study 6.

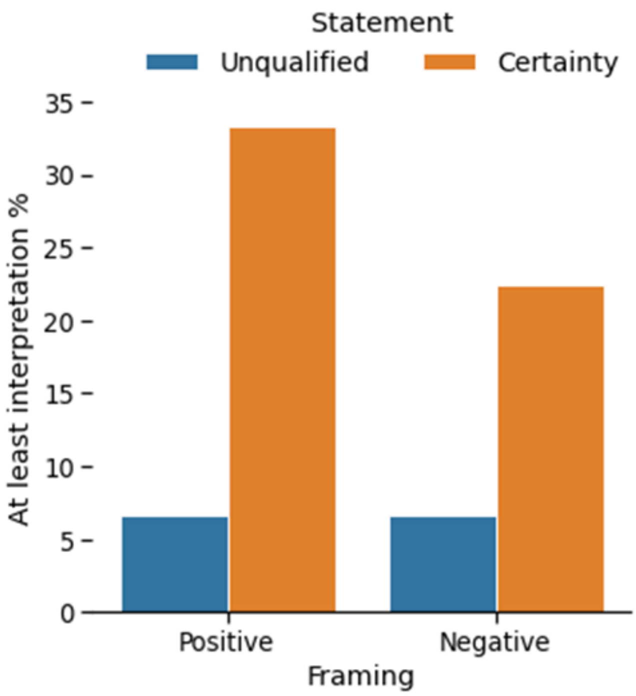

Figure 3. Percentages of *at least*-interpretations for plain (unqualified) vs. certainty-statements of program A in the positively framed condition (200-600 lives saved) and in the negatively framed condition (400-600 lives lost), Study 7.

8.3. Discussion

Unqualified statements might be judged as more concrete than statements that include surplus information about certainty. But concreteness can also refer to the statement's truth status, in which case objective impersonal certainty words like "definitely" and "guaranteed" can make the statement sound more settled and hence more concrete.

Judgments of Plan A in the risky choice scenario showed once again that certainty-statements are prone to lower-bound interpretations more often than plain, unqualified statements. But this could hardly count as a full explanation of the large framing effects, as even those who gave an "exact" interpretation of Plan A were not immune to framing. Thus the suggested lower-bound account for the framing effect (Mandel, 2014) was only partially supported by the present data.. Interestingly, the interpretation seem to depend on the valence of the outcome described, with desirable outcomes (lives saved) given at-least interpretations more often than aversive outcomes (lives lost).

9. Study 8

The preceding studies showed that unqualified quantitative statements by lay people tended to be more believable than statements claimed to be certain. However, there were exceptions. The climate experts in Study 2 were considered *more* trustworthy when expressing certainty in impersonal terms, whereas Study 3 indicated that they were *less* trustworthy when expressing personal certainty. Both these studies used a within-subjects design. A difference between personal and impersonal certainty was replicated in dialogues between non-experts reported in Study 5, using a between-subjects design. These dialogues also indicated that unqualified estimates were believed to rely more strongly on objective evidence than estimates claimed to be certain.

Studies 8 and 9 were designed to extend these findings to claims made by scientists. We must assume that scientists base their estimates on objective evidence to a greater extent than lay people engaged in everyday conversations do. Yet they are bound to make individual judgments based on this evidence, too, often leading to less than 100% consensus among their peers.

Scientific claims are sometimes reported in the media with headlines describing what "scientists" say or believe without identifying the specific source. Such generic statements imply some generality of findings but remain silent about how large is the consensus group. Haigh et al. (2020) asked people to estimate the degree of scientific consensus s implied by common generic phrases such as "Scientists say…" and "Experts believe…" and found wide individual variations, around a mean estimate of about 60%. In their study the phrases were presented without context, which may have added to the variability since judgments of consensus clearly will depend on the assertions made and how plausible they appear.

In contrast, the present studies described scientists' predictions about future climate changes, a theme most respondents is bound to recognize. To make the prospects slightly less familiar the predicted effects were to take place Africa and Australia. The predictions were attributed either to a singular (unnamed) scientist or to an (indeterminate) plural of "scientists". We assume that even the singular scientist will be regarded as a representative of the scientific community in some way or another, if nothing else by relying on scientific methods, so his predictions are probably not unique, but perhaps not shared by all his peers. Accordingly, we expected that predictions attributed to a plurality of “scientists” would be judged as more consensual than claims attributed to a singular “scientist”. More importantly, we expected that estimates where certainty is not mentioned will be rated as more strongly based on evidence than those claimed to be “certain”. Materials and data for this study is available on https://osf.io/2beh6/?view_only=fc7b9dc83c074655824516725bad4f20

9.1. Method

9.1.1. Participants

Participants were recruited by Prolific as in the previous studies. Of the 361 questionnaires collected, 19 were blank or incomplete, the remaining 342 were deemed valid and were retained for analysis (completion time > 1. min); 170 men and 170 women, 2 other or preferred not to say; mean age = 41.3 years (*SD* = 11.7); 94.2% identified themselves as native English speakers, and 58.5% held the equivalent of a bachelor degree or higher. Participants were randomly allocated to 2 x 2 conditions, with scientist (singular vs. plural) and type of claim (unqualified vs. personal certainty) as between-subjects factors.

9.1.2. Materials and procedure

After a brief, unrelated questionnaire, participants were introduced to two climate predictions as follows: “In the future, climate scientists expect higher mean temperatures all over the globe by the year 2100, along with more precipitation (rain) in wet areas and less precipitation in arid (dry) areas. We will present you two newspaper headlines related to climate change, and for each, we would like to know your perception of it.”

Participants in the *unqualified* conditions were presented with two climate predictions, formulated as newspaper headlines, presented in random order. [Plural condition in brackets]

Precipitation. Imagine the following headline in a newspaper you trust: “Climate scientist says [Climate scientists say] that in Africa, precipitation will decrease by 20% in the next decade”.

Temperature. Imagine the following headline in a newspaper you trust: “Climate scientist says [Climate scientists say] that by 2050, the mean temperature in Australia will be at least 1.5 degrees warmer than this year”.

In the *qualified* conditions the corresponding statements were: “Climate scientist says she is certain …” (singular) or “Climate scientists say they are certain …” (plural). In the temperature prediction, “she” was replaced by “he”.

Both predictions were followed by one question about evidence and one about consensus.

1. To what extent do you think the prediction is based on objective scientific evidence? To be rated by a slider on a 11 point scale from 0: not based on objective evidence at all, to 100: fully based on objective evidence.
2. If we present the prediction above to 100 climate scientists, how many do you think would agree with it? To be answered with a self-generated number.

9.3. Results

9.3.1. Consensus. Mean estimates of agreement among scientists, displayed in the upper panel of Table 9 show that plain predictions of precipitation were believed to generate higher consensus than those claimed to be certain, and predictions made by "scientists" (plural) reflected a higher consensus than those attributed to a singular scientist.

In the precipitation scenario, a 2 x 2 ANOVA showed significant main effects, , for unqualified vs. qualified and plural vs. singular, respectively, and no significant interactions $F(1, 338) = 6.22$, $p < .001$, and $F(1, 338) = 10.47$, $p < .001$, $F(1, 338) = 0.70$, $p = .403$. In the temperatures vignette, plural led again to higher consensus than singular, whereas the difference in consensus for unqualified vs. qualified statements was not significant, while there was no interaction, $F(1, 338) = 29.52$, $p < .001$, $F(1, 338) = 3.01$, $p =.084$, $F(1, 338) = 3.02$, $p = .083$.

Table 9. Mean estimates of consensus and objective evidence for unqualified (plain) vs. certain predictions of future precipitation and temperatures, made by a singular scientist or a generic plural of "scientists say", in Study 8 (Standard deviations in parentheses).

	Precipitation		Temperatures	
	Plain	Personal certainty	Plain	Personal certainty
Consensus				
Singular	70.2 (22.4)	61.7 (26.4)	73.3 (21.6)	65.5 (26.3)
Plural	76.4 (19.7)	72.1 (26.0)	81.6 (18.1)	81.6 (16.7)
Evidence				
Singular	69.7 (21.5)	59.6 (24.4)	70.0 (22.5)	65.7 (23.1)

Plural	73.5 (19.8)	66.7 (23.6)	79.0 (18.0)	73.8 (20.8)

9.3.2. Evidence. Mean estimates of evidence, displayed in the lower panel of Table 9, indicate that plain estimates were consistently judged to be more strongly based on objective scientific evidence. Unqualified estimates were judged to be significantly higher than certain estimates both for precipitation, $F(1, 338) = 12.17$, $p < .001$, and for temperatures, $F(1, 338) = 4.30$, $p = .039$. Again, estimates made by a plurality of "scientists" were judged to be more strongly based on objective evidence than those made by a singular scientist, $F(1, 338) = 5.06$, $p = .025$, and $F(1, 338) = 13.93$, $p < .001$, for precipitation and temperatures, respectively.

9.4. Discussion

The results indicate that plain forecasts are believed to rely more strongly upon evidence than personal certainty forecasts. Also, estimates made in unqualified statements were believed to be shared by more scientists than estimates that were certain. Agreement among scientists (consensus) about an estimate can be viewed as a measure of its believability, or trustworthiness.

Thus the results confirmed hypothesis 2b (Statements qualified by "certain" convey a *less* precise and reliable estimate than plain statements) and were in line with findings from the previous studies where participants appreciated unqualified statements above claims of personal certainty. This preference is apparently not limited to first person statements (I am certain) but can be extended to personal certainty of scientists using third person (he, she) singular or even plural (they).

10. Study 9

Results from the climate vignette in Study 2 and 3 taken together indicated that impersonal certainty increased confidence in predictions made by scientists, whereas personal certainty made the statements less believable. Study 8 confirmed the advantage of plain statements over personal certainty in a between-subjects design, and also that such statements were believed to rest more strongly up on evidence. The purpose of the present study was to compare unqualified estimates with both kinds of certainty estimates in the same design. The study was preregistered. We expected plain statements and statements of impersonal certainty to be perceived as more based on scientific evidence and more likely to come true than statements of personal certainty. We also expected the impersonal certainty to be similar or more likely to be based on scientific evidence and to come true than plain statements. The preregistration, along with materials and data is available on the OSF: https://osf.io/2beh6/?view_only=fc7b9dc83c074655824516725bad4f20.

10.1. Method

10.1.1. Participants

Participants were recruited by Prolific as in the previous studies. After 1 incomplete questionnaire was discarded, all of the 481 responses were deemed valid (completion time > 1.5 min); 239 men and 235 women, 7 other or preferred not to say; mean age = 42.8 years (*SD* = 13.7); 94.0% identified themselves as native English speakers, and 57.0% held the equivalent of a bachelor degree or higher. Participants were randomly allocated to 3 conditions, with type of claim (unqualified vs. personal vs. impersonal certainty) as a between-subjects factor.

10.1.2. Materials and procedure

All participants received the same two climate forecasts as in Study 8, one concerning future precipitation in Africa and the other about future temperatures in Australia.

In the *unqualified* condition they were told: "A climate scientist who is asked to estimate future climate changes in Africa, says: «In Africa, precipitation will decrease by 20% in the next decade."

In the *personal certainty* condition, the climate scientist said: «I am certain that in Africa, precipitation will decrease by 20% in the next decade."

In the *impersonal certainty* condition, the climate scientist said: «It is certain that in Africa, precipitation will decrease by 20% in the next decade."

Participants in all conditions were subsequently asked:

(1) To what extent do you think the prediction is based on objective scientific evidence? To be rated on an 11-point slider scale from 0: Not at all based on objective evidence, to 10: Fully based on objective evidence.

(2) Based on this prediction, what is the likelihood of a 20% precipitation decrease occurring in the next decade in Africa? Please answer by providing a percentage value ranging from 0 to 100.

The second forecast (future temperatures in Australia) was presented and responded to in the same way. The order of the two forecasts were counterbalanced across participants.

10.2 Results and discussion

Mean ratings (displayed in Table 10) demonstrate a quite consistent pattern: Temperature predictions were considered more likely and also more strongly based on evidence than precipitation predictions. They were probably also more aligned with the participants' own hypotheses, as global warming is a more generally acknowledged fact than regional differences in precipitation.

Personal certainty yielded lower score on ratings of evidence as well as probability, whereas unqualified estimates and impersonal certainty were at the same level. Thus the difference between personal and impersonal certainty was replicated, while plain statements were in this study not superior to certainty statements of an impersonal kind.

Table 10. *Mean estimates of evidence (1-10) and occurrence probabilities (0-100) for plain and certain forecasts by a climate expert, Study 9 (Standard deviations in parentheses)*

	Unqualified estimates	Personal certainty	Impersonal certainty	Overall *F* (2, 478)	Pairwise comparisons
Evidence					
Precipitation	6.54 (2.36)	6.19 (2.34)	6.48 (2.29)	1.04, p = .356	*ns*
Temperature	7.36 (1.91)	6.91 (1.93)	7.27 (2.04)	1.96, p = .142	I am vs. it is, p = .050
Probability					
Precipitation	57.7 (26.8)	55.2 (29.1)	61.6 (31.3)	2.41, p = .091	I am vs. plain p = .038
Temperature	69.3 (24.8)	65.2 (27.7)	73.5 (27.0)	3.83, p = .022	I am vs. it is, p = .006

The results from this experiment indicate that impersonal claims of certainty for climate predictions were more persuasive than expressions of personal certainty, at the same level (or even higher) than unqualified estimates. We did not give information about the setting where the statements were produced in this study, except that they came as answer to a question about "future climate changes" in Africa and Australia. Since the future generally is fraught with uncertainty, to say about a prediction that "it is certain" may be more informative (less redundant) than to emphasize certainty about more mundane, knowable facts, like "Jack has (certainly) two cars."

11. General discussion

When people estimate quantities or what may happen in the future they can indicate whether they are certain or not certain. Alternatively, they can choose to say nothing about the epistemic status of their estimates. Past work on verbal phrases expressing degrees of uncertainty (unlikely, likely, almost certain) shows that *certain* is interpreted to mean close to a 100% probability of a statement being true or an outcome occurring top (Clark, 1990; Clarke et al., 1992; Hamm, 1991; Mandel & Irwin, 2021; Reagan et al.1989; Witteman & Renooij, 2003; Willems et al., 2020). While verbal probabilities have been extensively studied and compared, particularly with respect to their probabilistic meanings, we know no studies where such qualified estimates have been compared to plain statements where degrees of certainty or probability have *not* been mentioned. We investigated three aspects of how statements made with certainty and unqualified ones might differ. Drawing on psychological and linguistic research, we proposed that expressing certainty (vs. not mentioning certainty) might actually lead to lower credibility, be indicative of vaguer outcomes, and be more likely to rely on guesswork than on objective evidence.

Credibility. The present studies show that unqualified, straightforward statements are often *more* trustworthy than statements qualified as certain. Statements with explicit certainty qualifications convey less confidence, and are perceived to be based on subjective judgment. This is particularly conspicuous when the certainty is expressed in phrases with personal pronouns ("I am certain") but also noticeable for external, impersonal attributions of certainty ("it is certain") in ordinary conversations. Thus, speakers who claim that estimates are "certain" do not necessarily make them more credible, on the contrary it could make them appear more questionable. We can accordingly conclude that we find more support for hypothesis H1b than for H1a: Adding certainty to a statement does not make it more likely, with the possible exception of impersonal statements about the future, expressed by experts. This occurred when statements with and without certainty were displayed jointly (Study 2). In such cases some participants might infer that statements which did not mention certainty were actually *not* certain.

Preciseness. We also found evidence of relevance for the second hypothesis that adding certainty to a numeric estimate would change the interpretation of the outcome into a lower bound of a prediction interval. These findings supported H2b more consistently than H2a, indicating that plain statements differ from certainty-statements as suggesting a higher degree of precision.

Source. Finally, impersonal certainty (it is certain) was generally superior to personal certainty, in line with hypothesis H3, and was occasionally, judged equally or more credible than plain statements. This occurred in cases where certainty and plain statements were compared in a joint presentation format (Study 2) and predicted future events (Study 9) but not for trivial facts that were rated separately (Study 4). In addition, explicit certainty reveals that estimates are based on judgments rather than on first-hand evidence.

11.1. The merits of plain assertions

Unqualified straightforward assertions have several advantages. They are shorter and hence might require less processing efforts, and are often perceived as more concrete. They ring true, as long as no alternative estimates are suggested. People draw inferences and conclusions most easily from the possibilities lined up in front of them, disregarding alternatives that they are not exposed to—a tendency that Kahneman (2011) has dubbed the WYSIATI principle (What You See Is All There Is). People tend to accept a proposition if it is not contested (Gilbert, 1991).

An estimate claimed to be *certain* represents, in comparison, a larger mental load. It contains a double message, one about the magnitude of estimate or outcome, the second about its certainty (a meta-cognitive message). A parallel can perhaps be drawn to the psychology of affirmations vs. negations. In Horn's (1989) apt phrase, affirmations describe states of the world, whereas negations describe *statements* about the states of the world, and hence are more abstract and require more mental resources.

11.2. The functions of explicit claims of certainty

Explicit claims of certainty have obviously a purpose. For instance, they could come in response to direct questions about the validity of a surprising or novel assertion. Should we believe this estimate? Yes, it is certain. However, to evaluate the legitimacy of such claims would require a more comprehensive analysis of discourse in context than was feasible with the current experimental approach, including exchanges between parties in a conversation. A limitation of the present studies is that the context provided for statements and estimates was minimal, with rudimentary dialogues occurring in only one of them (Study 5). We might assume that explicit reassurances of certainty are most appropriate and considered most informative in areas where facts are not easily accessible or cannot be taken for granted, for instance on topics that are questioned or surrounded by uncertainty. It may be no accident that the vignettes where explicit claims of certainty were considered most appropriate in the present set of studies, involved predictions (future sea levels and the election vignette in Study 2, and the climate projections of Study 9). These estimates all refer to the future, which often is considered uncertain by default, and accordingly legitimately examined before they are trusted, in line with predictions from the theory of epistemic vigilance (Sperber et al., 2010).

Statements emphasizing certainty also pragmatically implicate that other estimates are conceivable (but wrong), or that the estimate needs to be backed up by the speaker (in the absence of observable evidence). For instance, in a context of prior uncertainty (the watches showing different hours, or the future being inherently uncertain) a certainty statement may be considered more informative and preferred to an unmarked statement.

The present studies do not allow us to draw conclusions about the relative importance of these accounts. A cognitive load (or fluency) interpretation could be explored by giving people similar tasks under conditions of busyness or time pressure. A consensus interpretation could be investigated by asking participants about the number of experts they assume will endorse the statement. Another approach to understand the connotations of a certainty statement could be to view it as response to a question. If certain is the answer, what is the question? According to the alerting hypothesis (Isberner et al., 2013), the presence of epistemic adverbs will alert listeners to direct their attention to the validity of the propositions asserted in the sentence. Given that unmarked sentences are the default case, both certainty and uncertainty markers may be assumed to exert this effect.

11.3. Lower-bound interpretations of certainty

A recurrent observation in the present studies is the more frequent lower-bound interpretation of certainty estimates. Estimates claimed to be certain were frequently given an at least-interpretation in Study 1-3, whereas unmarked statements were perceived more often as point estimates. This finding was foreshadowed in a sentence-completion study by Teigen et al. (2014), where participants were shown a complete outcome distribution of the duration of computer batteries and asked to fill in an appropriate estimate in statements about "certain" (vs. "possible") outcome values. As there was no single value in the distribution that occurred with a 100% frequency, a majority proposed duration estimates in the low end of the distribution. These were perceived as *certain* life time of batteries, presumably because all of them lasted for this amount of time *or more.* In the past study, participants had to complete certainty statements (what is *certain* to happen and what *will* happen) with an appropriate number, while in the present studies they were provided with full statements and asked what the numbers meant, and how credible they were.

A lower bound interpretation of certain implies that such estimates are imprecise, suggesting a one-sided interval estimate rather than a point value. A corollary of this lack of precision is that inaccurate certainty statements might be considered less wrong than inaccurate plain statements. This will primarily occur for underestimates, as shown by Study 5 and 6. A financial expert who predicts that stocks will "certainly" increase with 3% may be regarded as (partly) correct if their value increases with 4%, but not when they increase with 2%. An advisor who predicts an unqualified, plain increase of 3% may be considered wrong in both these cases. Hypothetical *at least*-interpretations of numbers have played a key role in debates about the irrationality of framing effects (Fisher & Mandel, 2021; Geurts, 2013). Our results add to this discussion by showing that lower bound interpretations are particularly widespread for certainty statements, like those used in Mandel's (2014) studies, and may facilitate, but not account for all instances of framing effects.

11.4. The augmentation hypothesis

The present studies do not explain *why* certain-statements appear lower-bounded and why plain statements are assumed to be more exact, believable, and objective, apart from suggesting that the latter may be more immediate and concrete and do not ask the listener to reflect upon the statements' epistemic status. But claims of certainty may also have an argumentative, pragmatic function, by "leaking" information (Sher & McKenzie, 2006) about potency and strength. A loss that is *certain* to occur is one that must be taken seriously, a *certain* death is one from which there is no escape. To emphasize agreement with a proposition, respondents do not just say "yes", but "yes, for sure" or "yes, absolutely". A speaker may say "certainly" to reinforce or augment a message over and above its likelihood. For instance, when people say: "I certainly have two hundred Euro", they want to tell you that they have an abundance of money, not limited to four 50-Euro bills. Previous studies of single bound estimates (Teigen 2008; Teigen, Halberg & Fostervold, 2007) have shown that lower bound expressions, like "more than" and "at least", are pragmatically interpreted to mean "a lot". They have an upward, positive directionality (Teigen, 2022). An athlete who is *certain* to win a race is expected to win it with a large margin. A concert ticket that will *certainly* cost 10 euro may turn out to cost 20 euro, rather than five euro or even 10 euro and 50 cent.

This "more-than", augmentative aspect of certainty-statements might be tested more directly in further studies. For instance, imagine that a climate expert says in a plain condition that sea level will rise with 60 cm by the year 2100. If you are a city planner, constructing dikes to keep the city safe, how much higher should the dikes be compared to those that are in use today (which have worked well till now). Participants in a certainty-condition are told that the expert says the sea level is *certain* to rise with 60 cm. We predict that the second city planner will propose higher dikes than the former, not because his probabilities are larger, but because the rise in sea level is assumed to be higher.

A related study could be modelled after the studies of Mislavsky and Gaertig (2021) and Teigen et al. (2023). In these studies, the probability estimates from two, independent sources were combined. The resultant estimate increased rather than being merely averaged when based on positive phrases (likely, a chance) and was lower than initial estimates expressed with negative phrases (unlikely, not certain). The strongly positive expression "certain" was not examined in these studies. We predict that two experts stating that an investment will *certainly* be profitable may reinforce each other so that listeners conclude that the combined chances of its future profitability are higher than when based on just a single forecast. But the current studies also suggest that certainty conveys, in addition to a high probability, a message about strength and amplitude. It follows that an analyst who claims that an investment is "certain" to be profitable might not just expect a small or moderate (e.g., a 2-3%) rise, but a sizable (e.g., a 5-10%) increase. If two analysts, independently of each other, claim that they are certain, the expected rise might be even larger.

The admittedly speculative augmentation hypothesis fits well with a dispositional or propensity account of uncertainty (Kahneman & Varey, 1990; Popper, 1959) where outcome magnitudes and probabilities are seen to be in correspondence with each other (Keren & Teigen, 2001). According to this account, probabilities can be assigned to singular events by identifying causal mechanisms. High probabilities will, in turn, be expected to have stronger and sooner effects than low probabilities (Løhre, 2018).

Such surplus meanings of certainty-statements may depend upon the conversational setting. We may assume that they primarily apply to situations where speakers choose freely what to say and deliberately include or omit a reference to certainty. For instance when asked about the time, it is up to Jill in Study 5 to mention or not mention her degree of certainty. In other contexts an estimate is already around (provided by the speaker or suggested by a third party), the question being its validity. In such a context "certainly" may indicate a stronger agreement than simply "yes". When statements are evaluated out of context, and people just are asked about their trust, they must imagine for themselves a likely setting.

11.5 Practical implications

The findings have practical implications for communicators, by showing that a claim of certainty may be less trustworthy than an assertion accompanied by no such claims. A doctor may appear more reassuring and believable by merely stating that the patient will recover, than by saying that recovery is "certain" to occur. A climate expert predicting a global warming of 2 degrees may sound more authoritative than one asserting that this temperature rise is "certain", but on the other hand, announcing that this temperature increase is certain may be more alarming by implying that the rise will be 2 degrees *or more*. By using the term certain about an estimate, scientists could be aware that it might be disputed, since "certainty" is associated with subjective judgments rather than with scientific facts.

The persuasiveness of unqualified statements does not invalidate the "confidence heuristic" (Price & Stone, 2004), which says that statements issued with certainty have better odds of being believed than uncertain statements. Rather, our findings imply that it is not necessary and perhaps not advisable to declare one's certainty aloud. Certainty is conveyed more unambiguously by factual statements where certainty is not mentioned, perhaps related to literary authors' adage: "Show—don't tell".

The current research has implications for standards of verbal phrases recommended by international organizations like the IPCC (Mastrandrea et al., 2010), EFSA (European Food Safety Authority et al., 2018) and NATO (Mandel & Irwin, 2021) where *certain* is placed at the top of the probability scale. These scales typically assume probabilities for dichotomous (binary) events, but are less well-chosen when used to characterize quantities and amounts (Teigen et al., 2022). For instance, a "likely" value in a distribution may indicate a central, representative outcome whose probability is far below the stipulated 60-80%. "Certain" estimates may correctly reflect a 100% confidence, but on the other hand with less exactitude than commonly assumed, and with the suggestion that the outcome is a minimum to be expected. And the translation standards are silent on the fact that unqualified estimates may be preferable when people ask for "facts".

References

Breheny, R. (2007). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics, 25*, 93–139. https://doi.org/10.1093/jos/ffm016

Chick, C. F., Reyna, V. F., & Corbin, J. C. (2016). Framing effects are robust to linguistic disambiguation: A critical test of contemporary theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(2), 238-256. https://doi.org/10.1037/xlm0000158.

Clark. D. A. (1990). Verbal uncertainty expressions: a critical review of two decades of research. *Current Psychology: Research & Reviews*, *9*(3), 203–235.

Clarke, V. A., Ruffin, C. L., Hill, D. J., & Beamen, A. L. (1992). Ratings of orally presented verbal expressions of probability by a heterogeneous sample. *Journal of Applied Social Psychology, 22*, 638-656. https://doi.org/10.1111/j.1559-1816.1992.tb00995.x

Collins, P. J., & Hahn, U. (2018). Communicating and reasoning with verbal probability expressions. *Psychology of Learning and Motivation, 69*, 67–105. https://doi.org/10.1016/bs.plm.2018. 10.003

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

European Food Safety Authority et al. (2018). Guidance on communication of uncertainty in scientifc assessments. EFSA Journal. https://doi.org/10.2903/j.efsa.2019.5520

Ferson, S., O'Rawe, J., Antonenko, A., Siegrist, J., Mickley, J., Luhmann, C. C., et al. (2015). Natural language of uncertainty: Numeric hedge words. *International Journal of Approximate Reasoning, 57,* 19–39. https://doi.org/10.1016/j.ijar.2014.11.003

Fisher, S. A., & Mandel, D. R. (2021). Risky-choice framing and rational decision-making. *Philosophy Compass, 16*, e12763. https://doi.org/ 10.1111/phc3.12763.

Fox, C. R., & Ülkümen, G. (2017). Comment on Løhre and Teigen. *Thinking & Reasoning*, *23* (4), 483-491. https/doi.org/10.1080/13546783.2017.1314939

Fox, C. R., & Ülkümen, G. (2011). Distinguishing two dimensions of uncertainty. In W. Brun, G. Keren, G. Kirkebøen, & H. Montgomery (Eds.), *Perspectives on thinking, judging, and decision making* (pp. 21–35). Oslo: Universitetsforlaget.

Frisch, D. (1993). Reasons for framing effects. Organizational Behavior and Human Decision Processes, 54, 399–429. https://doi.org/ 10.1006/obhd.1993.1017

Geurts, B. (2013). Alternatives in framing and decision making. *Mind & Language, 28,* 1–19. https://doi.org/10.1111/mila.12005

Geurts, B., & Nouwen, R. (2007). At least et al.: The semantics of scalar modifiers. *Language, 83*(3), 533-559. https://www.jstor.org/stable/40070901

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*(2), 107-119. https://doi.org/10.1037/0003-066X.46.2.107

Gillies, D. (2000). *Philosophical theories of probability.* London: Routledge. https://doi.org/10.4324/9780203132241

Grice, H. P. (1975). Logic and conversation. In R. Cole & J. Morgan (Eds.), *Syntax and semantics: speech acts*. Academic Press.

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

Hacking, I. (1975). *The emergence of probability: A philosophical study of early ideas about probability, introduction and statistical inference*. Cambridge: Cambridge University Press.

Hamm, R. M. (1991). Selection of verbal probabilities: A solution for some problems of verbal probability expression. *Organizational Behavior and Human Decision Processes, 48*(2), 193-223. https://doi.org/10.1016/0749-5978(91)90012-I

Hansen, J., & Wänke, M. (2010). Truth from language and truth from fit: the impact of linguistic concreteness and level of construal on subjective truth. *Personality and Social Psychology Bulletin, 36*(11), 1576–1588. https://doi.org/10.1177/0146167210386238

Horn, L. R. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Hsee, C. K. (1998). Less is better: When low-value options are judged more highly than high-value options. *Journal of Behavioral Decision Making, 11*, 107–121.

https://doi.org/10.1002/(SICI)1099-0771(199806)11:2%3C107::AID-BDM292%3E3.0.CO;2-Y

Hsee, C. K. & Zhang, J. (2004). Distinction bias: misprediction and mischoice due to joint evaluation. *Journal of Personality and Social Psychology, 86*, 680–695. https://psycnet.apa.org/doi/10.1037/0022-3514.86.5.680

Isberner, M.-B., Richter, T., & Kaakinen, J. (2013). Epistemic modality in sentence comprehension: Effects of epistemic adverbs on eye movements. In M.-B. Isberner, *The role of epistemic monitoring in language comprehension*. PhD Dissertation, University of Kassel.

Jerez-Fernandez, A., Angulo, A. N., & Oppenheimer, D. M. (2014). Show me the numbers: Precision as a cue to others' confidence. *Psychological Science, 25*(2), 633–635. https://doi.org/10.1177/0956797613504301

Juanchich, M., Teigen, K. H., & Gourdon, A. (2013). Top scores are possible, bottom scores are certain (and middle scores are not worth mentioning): A pragmatic view of verbal probabilities. *Judgment and Decision Making, 8*(3), 345—364. https://doi.org/10.1017/S193029750000601X

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*, 445–465. https://doi.org/10.1017/S1930297500006483

1. Juanchich, M. and M. Sirota, *Do people really prefer verbal probabilities?* Psychological research, 2020. **84**(8): p. 2325-2338.

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

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

Kahneman, D., & Varey, C.A. (1990). Propensities and counterfactuals: The loser that almost won. *Journal of Personality and Social Psychology. 59*(6). 1101-1110. https://psycnet.apa.org/doi/10.1037/0022-3514.59.6.1101

Keren, G., & Teigen, K. H. (2001). The probability-outcome correspondence principle: A dispositional view of the interpretation of probability statements. *Memory & Cognition, 29*, 1010-1021. https://doi.org/10.3758/BF03195763

Kühberger, A., & Tanner, C. (2010). Risky choice framing: Task versions and a comparison of prospect theory and fuzzy-trace theory. *Journal of Behavioral Decision Making, 23*, 314–329. https://doi.org/10.1002/bdm.656

LaCour, M., Serra, M. J., Duval, M., & Hislop, C. (2023). Getting lost in the ”realm of possibility”. Common phrases used to communicate rare events have substantially different effects on decision-making. *Journal of International Crisis and Risk Communication Research*, *6*(2), 30-52. https://doi.org/10.56801/jicrcr.V6.i2.3

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology, 33*(4), 378–392. https://doi.org/10.1177/0261927X14535916

Lewandowsky, S. (2020). The 'post-truth' world, misinformation, and information literacy: A perspective from cognitive science. In S. Goldstein (Ed.), *Informed societies: Why information literacy matters for citizenship, participation and democracy* (pp. 69-88). Facet Publishing.

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,* 2531-2547. https://doi.org/10.1177/1747021817746062

Løhre, E., Juanchich, M., Sirota, M., Teigen, K. H., & Shepherd, T. G. (2019). Climate scientists' wide prediction intervals may be more likely but are perceived to be less certain. *Weather, Climate and Society, 11*(3), 565–575. https://doi.org/10.1175/ WCAS-D-18-0136.1

Løhre, E., Chandrashekar, S. P., Mayiwar, L., & Hærem, T. (2024). Uncertainty, expertise, and persuasion: A replication and extension of Karmakar and Tormala (2010). *Journal of Experimental Social Psychology*, *113*. https://doi.org/10.1016/j.jesp.2024.104619.

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*, 369–396. https://doi.org/10.1080/13546783.2015.1069758

Løhre, E., & Teigen, K. H. (2017). Probabilities associated with precise and vague forecasts. *Journal of Behavioral Decision Making, 30*, 1014–1026. https://doi.org/10.1002/bdm.2021

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

Macdonald, R. R. (1986). Credible conceptions and implausible probabilities. *British Journal of Mathematical and Statistical Psychology, 39*, 15–27. https://doi.org/10.1111/j.2044-8317.1986.tb00842.

McKenzie, C. R. M., & Nelson, J. D. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing efects. *Psychonomic Bulletin and Review, 10,* 596–602. https:// doi.org/10.3758/BF03196520

Mandel, D. R. (2001). Gain-loss framing and choice: Separating outcome formulations from descriptor formulations. *Organizational Behavior and Human Decision Processes, 85*, 56 – 76. https://doi.org/10.1006/obhd.2000 .2932

Mandel, D. (2014). Do framing effects reveal irrational choice? *Journal of Experimental Psychology: General, 143*, 1185– 1198. https://doi.org/10.1037/a0034207

Mandel, D. R., & Irwin, D. (2021). Uncertainty, Intelligence, and National Security Decisionmaking, *International Journal of Intelligence and CounterIntelligence, 34*:3, 558-582. https://doi.org/ 10.1080/08850607.2020.1809056

Mastrandrea, M. D., Field, C. B., Stocker, T. F., Edenhofer, O., Ebi, K. L., Frame, D. J., …, Matschoss, P. R. (2010). Guidance note for lead authors of the IPCC ffth assessment report on consistent treatment of uncertainties. Intergovernmental Panel on Climate Change (IPCC).

Mislavsky, R., & Gaertig, C. (2021). 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

Okder, H. (2012). The illusion of the framing effect in risky decision making. *Journal of Behavioral Decision Making, 25*(1), 63–73. https://doi.org/10.1002/bdm.715.

Onay, S., La-ornual, D., & Öncüler, A. (2013). The effect of temporal distance on attitudes toward imprecise probabilities and imprecise outcomes. *Journal of Behavioral Decision Making, 26*, 362–374. https:// doi.org/10.1002/bdm.1763

Popper, K. (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science, 10*, 25-42.

Pezzuti, T., Leonhardt, J. M., & Warren, C. (2021). Certainty in language increases consumer engagement on social media. *Journal of Interactive Marketing, 53*, 32-46. https://doi.org/10.1016/j.intmar.2020.06.005

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

Reagan, R. T., Mosteller, F., & Youtz, C. (1989). Quantitative meanings of verbal probability expressions. *Journal of Applied Psychology, 74*(3), 433-442. https://doi.org/10.1037/0021-9010.74.3.433

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology, 96*(3), 538–558. https://psycnet.apa.org/doi/10.1037/a0014038

Scott, S. E., & Rozin, P. (2017). Are additives unnatural? Generality and mechanisms of additivity dominance. *Judgment and Decision Making, 12*(6), 572–583. https://doi.org/10.1017/S1930297500006707

Sher, S., & McKenzie, C. R. M. (2006). Information leakage from logically equivalent frames. *Cognition, 101*(3), 467–494. https://doi.org/10.1016/j.cognition.2005.11.001

Spector, B. (2013). Bare numerals and scalar implicatures. *Language and Linguistics Compass, 7*(5), 273–294. https://doi.org/10.1111/lnc3.12018

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & language*, *25*(4), 359-393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Steiger, A., & Kühberger, A. (2018). A meta-analytic re-appraisal of the framing effect. *Zeitschrift für Psychologie, 226*(1), 45–55. https://doi.org.eproxy.uio.no/10.1027/2151-2604/a000321

Teigen, K. H. (2008). More than X is a lot: Pragmatic implicatures of one-sided intervals. *Social Cognition, 26,* 379–400. https://doi.org/10.1521/soco.2008.26.4.379

Teigen, K. H. (2022). Dimensions of uncertainty communication: What is conveyed by verbal terms and numeric range estimates. *Current Psychology.* https://doi.org/ 10.1007/s12144-022-03985-0

Teigen, K. H., & Filkukova , P. (2011). Are lies more wrong than errors? Accuracy judgments of inaccurate statements. *Scandinavian Journal of Psychology, 52,* 8-20. https://doi.org/ 10.1111/j.1467-9450.2010.00843.x

Teigen, K. H., & Filkuková, P. (2013). Can > will: Predictions of what can happen are extreme, but believed to be probable. *Journal of Behavioral Decision Making, 26*(1), 68-78. https://doi.org/10.1002/bdm.761

Teigen, K. H., Halberg, A. M., & Fostervold, K. I. (2007). Single-limit interval estimates as reference points. *Applied Cognitive Psychology, 21,* 383–406. https://doi.org/10.1002/acp.1283

Teigen, K. H., Juanchich, M., & Filkuková, P. (2014). Verbal probabilities: An alternative approach. *Quarterly Journal of Experimental Psychology, 67,* 124-146. 1 https://doi.org/0.1080/17470218.2013.793731

Teigen, K. H., Juanchich, M., & Løhre, E. (2022). What is a "likely" amount? Representative (modal) values are considered likely even when their probabilities are low. *Organizational Behavior and Human Decision Processes*, *177*. https://doi.org/10.1016/j.obhdp.2022.104166

Teigen, K. H., Juanchich, M., & Løhre, E. (2023). Combining verbal forecasts: The role of directionality and the reinforcement effect. *Journal of Behavioral Decision Making, 36*(2), e2298. https://doi.org/10.1002/bdm.2298

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

Teigen, K. H. & Nikolaisen, M. I. (2009). Incorrect estimates and false reports: How framing modifies truth. *Thinking & Reasoning, 15*, 268–293. https://doi.org/10.1080/13546780903020999

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. https://psycnet.apa.org/doi/10.1037/a0018963

Trope, Y., Liberman, N., & Wakslak C. J. (2007). Construal levels and psychological distance: Effects on representation, prediction, evaluation, and behavior. *Journal of Consumer Psychology, 17*, 83–95. https://doi.org/10.1016/S1057-7408(07)70013-X

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453– 458. https://doi.org/ 10.1126/science.7455683

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

Van der Bles, A. M., van der Linden, S., Freeman, A. L. J., Mitchell, J., Galvao, A. B., Zaval, L., & Spiegelhalter, D. J. (2019). Communicating uncertainty about facts, numbers and science. *Royal Society Open Science, 6*, 181870. https://doi.org/10.1098/rsos.181870

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*, 7672-7683. https://doi.org/10.1073/pnas.1913678117

Vogel, H., Appelbaum, S., Haller, H., & Ostermann, T. (2022) The interpretation of verbal probabilities: A systematic literature review and meta-analysis. In R. Röhrig et al. (Eds.), *General Medical Data Sciences – Future Medicine* (pp. 9-16). IOS Press. https://doi.org/10.3233/SHTI220798

Wakslak, C. J., & Trope, Y. (2009). The effect of construal level on subjective probability estimates. *Psychological Science, 20*, 52-58. https://doi.org/10.1111/j.1467-9280.2008.02250.x

Wallsten, T. S., Budescu, D. V., Rapoport, A., Zwick, R., & Forsyth, B. (1986). Measuring the vague meanings of probability terms. *Journal of Experimental Psychology: General, 115*(4), 348-365.

Willems, S., Albers, C., & Smeets, I. (2020). Variability in the interpretation of probability phrases used in Dutch news articles-a risk for miscommunication. *Journal of Science Communication*, *19*(2). https://doi.org/10.22323/2.19020203

Wintle, B. C., Fraser, H., Wills, B. C., Nicholson, A. E., & Fidler, F. (2019). Verbal probabilities: Very likely to be somewhat more confusing than numbers. *PLoS ONE, 14*(4), e0213522. https:// doi.org/10.1371/journal.pone.0213522

Witteman, C., & Renooij, S. (2003). Evaluation of a verbal-numerical probability scale. International Journal of Approximate Reasoning, 33(2). https://doi.org/ 10.1016/S0888-613X(02)00151-2