Modified numerals:
Two routes to ignorance

Alexandre Cremers, Liz Coppock, Jakub Dotlacil & Floris Roelofsen

November 18, 2017

Abstract
Modified numerals, such as at least three or more than five, are known to sometimes give rise to ignorance inferences, in that their use suggests that the speaker does not have exact knowledge. However, there have been disagreements regarding the nature of these inferences, their robustness and context dependency, and differences between at least and more than. We first present a series of experiments which sheds new light on these issues and explains some contradicting reports in the previous literature. Our results show that (a) the ignorance inferences of at least tend to be more robust than those of more than, but (b) both modifiers are sensitive to the question under discussion (QUD). We argue that this pattern can be explained if we assume two sources of ignorance: traditional Quantity implicatures, which arise regardless of the modifier but are sensitive to the QUD, and implicatures arising from the maxim of inquisitive sincerity, which are QUD-insensitive but only appear with at least.

1 Introduction
Modified numerals like at least n and more than n contrast in interesting ways with each other and with bare numerals in the implicatures that they give rise to. The basic empirical picture assumed in most work on the topic is as follows:

\[(1) \quad \begin{align*}
    a. & \quad \text{The house has } \textbf{three} \text{ bedrooms.} & \sim & \text{exactly three} & \sim \text{ignorance} \\
    b. & \quad \text{The house has } \textbf{at least three} \text{ bedrooms.} & \not\sim & \text{exactly three} & \not\sim \text{ignorance} \\
    c. & \quad \text{The house has } \textbf{more than two} \text{ bedrooms.} & \not\sim & \text{exactly three} & \not\sim \text{ignorance}
\end{align*}\]

At least three and more than two do not trigger an upper-bounding implicature, in the sense that they do not give rise to the inference that the house has exactly three bedrooms, in contrast with the bare numeral three. Moreover, it is usually assumed that at least triggers an ignorance implicature, unlike more than or bare numerals.

While the contrast in upper-bounding implicatures is uncontroversial, the ignorance implicatures triggered by modified numerals are subject to ongoing debate. Do superlative modifiers always convey ignorance? Do comparative modifiers never do so? And, moving from empirical issues to theoretical ones, how do these ignorance implicatures come about? What are the crucial pragmatic reasoning processes involved, and what are the basic semantic structures associated with superlative and comparative modifiers that feed these
processes? In order to delineate the contribution of the present paper, let us look at each of these issues in some more detail.

**The empirical issue: which modifiers convey ignorance, and when?**

Consider the following scenario. The emergency department of a certain hospital is required to have three physicians present at all times. Following a complaint by a patient who had to wait for several hours on Tuesday last week, police officers are investigating whether the requirement was satisfied. An employee of the hospital tells the officers:

(2) a. There were at least three physicians last Tuesday.
   b. There were more than two physicians last Tuesday.

Can the officers conclude that the employee doesn’t know exactly how many physicians there were? The received view is that the superlative modifier *at least* in (2a) does imply such ignorance, while the comparative modifier *more than* in (2b) doesn’t (see, e.g., Geurts and Nouwen, 2007; Büring, 2008; Cummins and Katsos, 2010; Nouwen, 2010; Kennedy, 2015). This assumption, however, is not universally shared. For instance, Westera and Brasoveanu (2014) (henceforth W&B) hold that both *at least* and *more than* can in principle generate ignorance implicatures; whether they do is determined by the question under discussion (QUD). More specifically, W&B propose that in response to a *how many* question like (3a), both *at least* and *more than* imply ignorance, while in response to a *polar* question like (3b), neither does.

(3) a. How many physicians were there last Tuesday? $\sim$ ignorance
   b. Were there enough physicians last Tuesday? $\not\sim$ ignorance

Mayr and Meyer (2014) (henceforth M&M) also propose that ignorance implicatures are sensitive to the QUD. However, the specific empirical generalization that they suggest is different from that of W&B. According to them, both types of modified numerals convey ignorance in response to *how many* questions such as (3a), while only *at least* conveys ignorance in response to polar questions such as (3b).

A third account where ignorance implicatures are sensitive to the QUD has been proposed by Coppock and Brochhagen (2013b) (henceforth C&B). However, the specific predictions of this account again differ from those of W&B and those of M&M. Namely, on the account of C&B neither *at least* nor *more than* conveys ignorance in response to polar questions, while only *at least* conveys ignorance in response to *how many* questions. Note that this is in a sense the reverse image of M&M’s view: QUD sensitivity is not assumed for *more than* but rather for *at least*.

The different views are summarized in Table 1. The fact that different authors have disagreed on some of the basic data shows the need for further experimental investigation. The present paper offers experimental data lending support to some of W&B’s, M&M’s, and C&B’s qualms with the standard empirical picture, but we also identify new contrasts. In particular, we find a three-way contrast in strength/robustness of the ignorance implicatures triggered by modified numerals:

1. The ignorance implicatures of *at least* in *how many* contexts are strong and robust: they are observed across different experimental settings.

2. The ignorance implicatures of (a) *at least* in polar contexts and (b) *more than* in *how many* contexts are less robust: they are not detected in all experimental settings.
Moreover, the experimental settings in which the former are detected are not the same as those in which the latter are detected. Finally, ignorance inferences are usually weaker in these cases (and certainly never stronger) than those of at least in how many contexts.

3. We did not find any evidence for ignorance implicatures with more than in polar contexts, suggesting that if there are any they must be very weak.

This three-way contrast has not previously been identified, and forms a challenge for previous accounts, as will be discussed below.

The theoretical issue: how do ignorance inferences arise?

There is a wide consensus in the literature that the ignorance conveyed by modified numerals is not directly encoded in their lexical semantics, but rather derived pragmatically (pace Geurts and Nouwen, 2007). The precise nature of the pragmatic reasoning that leads to these ignorance inferences, however, is a matter of ongoing debate.

One approach is to derive ignorance inferences from Grice’s maxim of quantity (Schwarz, 2013, 2016b; Mayr and Meyer, 2014; Westera and Brasoveanu, 2014; Kennedy, 2015, among others). On this approach, one way to account for differences between the various kinds of modified numerals is to assume that they have different formal pragmatic alternatives, which, depending on the QUD, may or may not play an active role in the derivation of quantity implicatures.

Another approach is to derive ignorance inferences from the maxim of inquisitive sincerity, which is not concerned with the informative content of the uttered sentence, but rather with its inquisitive content, i.e., the semantic alternatives that it introduces (Coppock and Brochhagen, 2013b). Differences between the various kinds of modified numerals are accounted for on this approach by assuming that they introduce different semantic alternatives.

The present paper argues for a dual route hypothesis: both Gricean quantity and inquisitive sincerity potentially play a role in the derivation of ignorance inferences. However, whether these maxims come into play depends on the construction and context at hand. In particular, Gricean quantity gives rise to ignorance inferences when formal alternatives are activated, which we will argue is the case in typical how many contexts but not in polar contexts, while inquisitive sincerity gives rise to ignorance inferences whenever semantic alternatives are activated, which we propose to be the case with at least but not with more than. Consequently, as depicted in Table 2, ignorance inferences triggered by at least in how many contexts are expected to be strong and robust across different experimental settings, while those triggered by at least in polar contexts or by more than in how many contexts are expected to be weaker and less robust, because they rely either on inquisitive sincerity or on Gricean quantity alone. Moreover, since the ignorance inferences triggered by at least in polar contexts and those triggered by more than in how many contexts are

<table>
<thead>
<tr>
<th>Modifier</th>
<th>QUD</th>
<th>Received view</th>
<th>W&amp;B</th>
<th>M&amp;M</th>
<th>C&amp;B</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least</td>
<td>How many</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Polar</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>More than</td>
<td>How many</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Polar</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 1: Empirical views on the ignorance implications of modified numerals.
Table 2: Source and strength/robustness of ignorance on the dual route hypothesis.

<table>
<thead>
<tr>
<th>Modifier</th>
<th>QUD</th>
<th>Source</th>
<th>Strength/robustness</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least</td>
<td>How many</td>
<td>Both sincerity and quantity</td>
<td>Strong</td>
</tr>
<tr>
<td></td>
<td>Polar</td>
<td>Only sincerity</td>
<td>Intermediate</td>
</tr>
<tr>
<td>More than</td>
<td>How many</td>
<td>Only quantity</td>
<td>Intermediate</td>
</tr>
<tr>
<td></td>
<td>Polar</td>
<td>None</td>
<td>Null</td>
</tr>
</tbody>
</table>

of a different nature under the present hypothesis, it is expected that they may differ from each other both in strength and in the kinds of experimental settings in which they are detectable. Finally, no ignorance inferences are expected to arise with more than in polar contexts, since neither inquisitive sincerity nor Gricean quantity is effective in this case.

Thus, we argue that the two most prominent approaches to explaining the ignorance inferences of modified numerals—one relying on Gricean quantity and the other on inquisitive sincerity—should not be regarded as being in competition, but rather complement each other and are both needed to capture the subtle empirical contrasts observed.

The structure of the paper is as follows. First, Section 2 provides more detailed discussion of quantity-based approaches, and Section 3 reviews the inquisitive sincerity-based approach. Against this background, our experimental work is presented in Section 4, our dual route account is developed in Section 5, and the predictions of this account are discussed in detail in Section 6, with particular reference to the experimental findings presented in Section 4. Finally, Section 7 concludes.

## 2 Approaches based on Gricean quantity

### 2.1 Schwarz

We will first review the quantity-based account of Schwarz (2013, 2016b).\(^1\) This account is only concerned with at least, leaving more than out of consideration.\(^2\) In Section 2.2 we will turn to the proposal of Mayr and Meyer (2014), which can be seen as an extension of Schwarz’s account to cover both at least and more than.

Semantically, Schwarz assumes that a sentence like (4) simply expresses the proposition consisting of all possible worlds where Sam hired two or more cooks.

\[(4) \text{ Sam hired at least two cooks.} \]

Further, Schwarz assumes that the formal alternatives for a sentence like (4) are determined by two interacting Horn scales: one involving numerals, \(\langle 1, 2, 3, \ldots \rangle\), and the other involving the modifier at least, which according to Schwarz forms a Horn scale with only. Thus for any natural number \(n\), both ‘Sam hired only \(n\) cooks’ and ‘Sam hired at least \(n\) cooks’

\(^1\)We will reformulate the account somewhat, in a way that will allow for easy comparison with our own proposal. The essence of the account, as well as its empirical predictions, will of course be maintained.

\(^2\)See Mayr (2013), Kennedy (2015), and Alrenga (2016) for closely related quantity-based proposals. The latter focuses just on at least, like Schwarz, but the former two explicitly treat more than as well. Kennedy (2015) predicts that more than does not yield ignorance implicatures, while at least does. Mayr (2013) is mainly concerned with upper-bounding implicatures and does not explicitly discuss the predictions of his analysis concerning ignorance implicatures; as far as we can tell, however, neither at least nor more than is predicted to trigger any ignorance implicatures on this account.
are formal alternatives for (4). Schwarz uses the following notation to compactly list these alternatives.

(5) Formal alternatives for (4):

\[
\begin{array}{cccc}
1 & 1 & 2 & 3 & 4 & \ldots \\
2 & 2 & 3 & 4 & \ldots \\
3 & 3 & 4 & \ldots \\
4 & 4 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\end{array}
\]

To articulate how Schwarz takes these formal alternatives to factor into pragmatic reasoning, we introduce the following background notions and notation. A speaker’s information state is modeled as a non-empty set of worlds. We say that an information state \(s\) supports a sentence \(\varphi\) iff \(s \subseteq \llbracket \varphi \rrbracket\). On the other hand, we say that \(s\) rejects a sentence \(\varphi\) iff \(s \cap \llbracket \varphi \rrbracket = \emptyset\). Finally, we use \(A_\varphi\) to denote the set of lexically determined formal alternatives for \(\varphi\).

Implicatures can be seen as imposing constraints on what the speaker’s information state might be. On Schwarz’s approach, they are derived using the following recipe. First, the quality implicature that the speaker’s information state \(s\) supports \(\varphi\) is derived:

(6) \(\text{quality}(\varphi) := \{s \mid s \text{ supports } \varphi\}\)

Then primary quantity implicatures are derived based on the assumption that the speaker’s state does not support any formal alternative \(\psi \in A_\varphi\) that is stronger than \(\varphi\) itself. We use \(A^<_\varphi\) to denote the set of those stronger alternatives:

(7) \(A^<_\varphi := \{\psi \in A_\varphi \mid [\psi] \subset [\varphi]\}\)

(8) \(\text{quantity}1(\varphi) := \{s \mid s \text{ does not support any } \psi \in A^<_\varphi\}\)

Finally, secondary quantity implicatures are derived by identifying all alternatives \(\psi \in A^<_\varphi\) that are innocently excludable w.r.t. \(\varphi\) (Gazdar, 1979; Fox, 2007). Informally speaking, \(\psi\) is innocently excludable if, whenever a set of alternatives in \(A^<_\varphi\) has been consistently rejected we can always go on to reject \(\psi\) in addition, maintaining consistency. More precisely: for every subset \(A'\) of \(A^<_\varphi\), if it is possible to find an information state in \(\text{quality}(\varphi) \cap \text{quantity}1(\varphi)\) that rejects every sentence in \(A'\), then it is also possible to find a state in \(\text{quality}(\varphi) \cap \text{quantity}1(\varphi)\) that rejects every sentence in \(A'\) as well as \(\psi\).

Secondary quantity implicatures are based on the assumption that the speaker’s information state rejects all innocently excludable alternatives in \(A^<_\varphi\).

(9) \(\text{quantity}2(\varphi) := \{s \mid s \text{ rejects every } \psi \in A^<_\varphi \text{ that is innocently excludable w.r.t. } \varphi\}\)

If a speaker with information state \(s\) utters a sentence \(\varphi\) in a cooperative conversation, it is assumed that \(s \in \text{quality}(\varphi) \cap \text{quantity}1(\varphi) \cap \text{quantity}2(\varphi)\).

None of the formal alternatives that Schwarz assumes for example (4) is innocently excludable. For instance, the assumption that the speaker’s information state \(s\) rejects [2] is incompatible with the assumption that it rejects [3,...], given the quality implicature that \(s\) supports [2,...]. So we get primary quantity implicatures, but no secondary ones. The primary quantity implicatures say that the speaker’s information state does not reject any of the formal alternatives that are stronger than [2,...]. Thus, ignorance is derived. On the other hand, it is predicted that the sentence does not have an upper-bounding implicature,
which in the case of bare numerals comes about as a secondary quantity implicature.

### 2.2 Mayr and Meyer

Mayr and Meyer (2014) (M&M) present a theory that is much in the same spirit as that of Schwarz (2016b), but covers comparative modifiers as well. More specifically, it is designed to account for the following empirical generalizations:

- Superlative modifiers never give rise to upper-bounding implicatures, but **always** yield ignorance implicatures;
- Comparative modifiers never give rise to upper-bounding implicatures, but **sometimes** do yield ignorance implicatures, depending on the question under discussion.

The first generalization is standard. The second generalization, however, diverges from what is typically assumed. To motivate it, M&M submit that *more than* and *at least* equally convey ignorance on the part of the speaker in the following pair of examples:

\[(10)\]

a. Q: What’s the distance between Ramallah and Jerusalem?
   A: It’s more than 10km. \(~\) ignorance

b. Q: What’s the distance between Ramallah and Jerusalem?
   A: It’s at least 10km. \(~\) ignorance

Moreover, to justify the claim that the question under discussion affects whether *more than* conveys ignorance, M&M submit that there is a contrast between (11a), which involves a *how many* question, and (11b), which involves a polar question.

\[(11)\]

a. Q: How many kids do you have?
   A: #I have more than three kids. \(~\) ignorance

b. Q: I need to double-check if you qualify for these benefits. You have three kids?
   A: I have more than three kids. \(\not\sim\) ignorance

Observe that a version of (11b) with *at least* in place of *more than*, given in (12b), is quite odd, in line with the assumption that *at least* always gives rise to ignorance implicatures, even in cases where *more than* does not.

\[(12)\]

a. Q: How many kids do you have?
   A: #I have at least three kids. \(~\) ignorance

b. Q: I need to double-check if you qualify for these benefits. You have three kids?
   A: #I have at least three kids. \(\not\sim\) ignorance

M&M’s theory differs from Schwarz’s in a number of respects of varying importance for the issue at hand. One difference is that they assume that the LF of a sentence may or may not include an exhaustivity operator EXH (following, e.g., Chierchia et al., 2012). This operator has the effect of negating any stronger formal alternatives that are innocently excludable. M&M assume a pragmatic principle requiring, roughly, that an interpretation involving EXH is to be preferred whenever it is stronger than the one without EXH, and relevant with respect to the question under discussion. This assumption obviates the need to compute secondary quantity implicatures, so that part of the account of Schwarz
presented above is not included in their theory. Primary quantity implicatures are derived in a similar way, however: if a speaker utters a sentence $\varphi$, then for every formal alternative $\psi$ of $\varphi$ that is relevant w.r.t. the question under discussion and stronger than $\varphi$, it is inferred that the speaker’s information state does not support $\psi$.

A second difference between M&M’s theory and Schwarz’s is that M&M make use of a general recipe for computing formal alternatives, adopted from Katzir (2007) and Fox and Katzir (2011), rather than stipulating them as Schwarz does. For $at least$, this recipe yields essentially the same result as Schwarz’s stipulations: the pertinent formal alternatives to $at least three$, for example, are of the form $at least n$ and $n$ (the latter of which are understood under an ‘exactly’ interpretation). For $more than three$, the pertinent alternatives are of the form $more than n$ and $n$.

Thirdly, and most importantly, M&M allow for the ‘pruning’ of formal alternatives, subject to a certain constraint (following Fox and Katzir, 2011). According to M&M, this constraint prevents the pruning of alternatives for $at least$ under any circumstance, and because the alternatives are still in place, ignorance implicatures are always generated just as under Schwarz’s theory. But with $more than$, the alternatives can be pruned. The constraint is as follows:

\begin{equation}
\text{(13) Constraint on Alternative Pruning} \quad \text{(M&M slide 29, Fox and Katzir, 2011)}
\end{equation}

Given a structure $\varphi$, the set $\text{ALT}(\varphi)$ can be pruned to a subset $A \subseteq \text{ALT}(\varphi)$ only if there is no distinct alternative $\psi \in \text{ALT}(\varphi)$ such that $\lfloor \text{EXH } A \varphi \rfloor = \lfloor \psi \rfloor$.

For example, for (11b), M&M consider the pruned set of alternatives $A = \{three, more than three\}$. If the sentence is exhaustified with respect to this set of alternatives $A$, the result is not equivalent to any of the original formal alternatives. More specifically, since neither of the alternatives in $A$ is stronger than the uttered sentence itself, exhaustification has no effect at all. Hence the constraint on alternative pruning is satisfied. Thus, given that primary quantity implicatures are now computed based on the pruned set of alternatives $A$, no ignorance implicature is predicted to arise.

M&M also consider the case of (12b) and the potential pruning $A = \{three, at least three\}$. Exhaustifying with respect to this pruned set of alternatives $A$ would yield a meaning equivalent to $at least four$, which is one of the original alternatives. This is because $three$ is a stronger and innocently excludable alternative that can therefore be targeted by EXH. Because this pruning violates the condition on alternative pruning, the alternatives stay as they are, giving rise to the ignorance implicature, just as under Schwarz’s account.

One question that is left open in M&M’s proposal is which potential prunings should be considered exactly. Example (12b), where the considered pruned set of alternatives is $\{three, at least three\}$, suggests that it is not just alternatives that are strictly relevant to the QUD that may remain after pruning. Depending on how the QUD is interpreted in the given example, either $three$ or $at least three$ is not directly relevant to it. This raises the question what would happen if the pruned set were just, say, $\{five\}$. The single alternative $five$ would be innocently excludable, and the resulting meaning, with a ‘gap’ for five as it were, would not be equivalent to any of the original alternatives. Thus it seems that this should be a licit pruning, yielding a very strange interpretation. Presumably the account should be amended so that the pruned sets consist only of QUD-relevant alternatives. As far as we can see, this would deliver the desired prediction for example (12b).

However, the following variant of (12b), which is actually more semantically parallel to (11b), appears to be a cause for concern:

\begin{equation}
\text{(14) Q: I need to double-check if you qualify for these benefits. You have three kids?}
\end{equation}
A: I have at least four kids.

If \{three, at least four\} is a licit pruning here, parallel to the case with more than, then ignorance implicatures are predicted to disappear in this case, contrary to M&M’s empirical claim that at least always yields ignorance implicatures, and in conflict with the intuition that the answer in (14) conveys just as much ignorance, and is therefore just as odd, as the answer in M&M’s example (12b).

We leave open whether this challenge can be addressed. It may be possible to derive the predictions that Mayr and Meyer (2014) intend to make within the confines of a quantity-based approach. We note, however, that the strategy of deriving both scalar and ignorance implicatures from exhaustivity-based reasoning about formal alternatives entails a very tight coupling between ignorance implicatures and upper-bounding implicatures. We suspect that this bond may be difficult to break, something that would be needed in order to account for the claimed difference in ignorance implicatures between at least and more than in the context of a polar QUD.

2.3 Westera and Brasoveanu

Westera and Brasoveanu (2014) (W&B), independently of M&M, also argue that the standard empirical assumptions about the ignorance inferences triggered by at least and more than are too simplistic. However, the empirical generalizations that they propose, driven
Figure 2: Westera and Brasoveanu’s (2014) results.

by experimental data, are different from those of M&M.

As depicted in Figure 1, W&B presented experimental participants with courtroom dialogues such as the one in (15):

(15) Judge: What did you see under the bed?
Witness: I saw at most 10 diamonds under the bed.

The type of question was experimentally manipulated as indicated in Figure 1, and the witness’s response always contained either a superlative modified numeral, *at most 10*, or a comparative modified numeral, *less than 10*. Participants were then told that the judge concluded that the witness does not know exactly how many diamonds she saw under the bed, and were asked how justified the judge was in drawing that conclusion, on a 1–5 scale.

From the results, given in Figure 2, W&B conclude that, contra the received view but in line with M&M, comparative modifiers sometimes *do* signal ignorance (e.g., in response to *how many exactly* questions), and also, contra the received view as well as M&M, that superlative modifiers sometimes *don’t* convey ignorance (most clearly in polar contexts). Extrapolating from these observations, W&B propose the following:

- When sentences with modified numerals are considered in the context of a fully explicit QUD, ignorance inferences are triggered if and only if the QUD requires an exact answer (e.g., in *how many exactly* context but not in *how many approximately* or polar contexts). This holds for superlative and comparative modified numerals alike, and is a result of straightforward quantity-based pragmatic reasoning: the speaker did not provide an exact answer, so, assuming that she intends to comply with the maxim of quantity, she must be unable to give an exact answer, i.e., she must be (partially) ignorant.

- If there is no explicit QUD, or if the QUD does not fully determine whether an exact answer is required or not (W&B assume that this is the case, for instance, for *how many* questions), the strength of the ignorance inference that arises depends on how likely it is that an exact answer is required in the context at hand. According to
W&B, this likelihood is higher with superlative modified numerals than with comparative ones, since the former are more often used in contexts where exact numerical information is required than the latter (they provide partial support for the latter claim based on a corpus-based study). Thus, it is predicted that in underspecified or ‘out of the blue’ contexts, at least triggers stronger ignorance inferences than more than.

We will make four brief remarks about W&B’s design, empirical results, and theoretical proposals at this point—a more in-depth discussion will be provided once our own experimental data have been presented. First, among the six types of context that W&B considered, a significant difference between superlative and comparative modifiers was only found in how many contexts. The fact that superlative and comparative modifiers generally behaved alike is in line with W&B’s view that, whenever the QUD is fully specified, there should be no difference between the two types of modifier. By contrast, it is surprising on the received view, on which superlative modifiers convey ignorance and comparative ones do not. This said, note that the fact that no significant difference was found between the two types of modifier (in most contexts) is a null result, and may be due in part to the experimental setting. That is, it may well be that in other experimental settings, differences between superlative and comparative modifiers can more easily be detected. This will indeed be the case in some of the experiments presented below.

Second, W&B’s explanation for the difference between superlative and comparative modifiers in how many contexts does not seem quite satisfactory, since it relies on the assumption that it is not clear in such contexts whether an exact answer is required. This assumption seems implausible: if a witness appears in front of a judge and is asked how many diamonds she saw under the bed, then she is clearly expected to say exactly how many diamonds she saw. An approximate answer is only in order if she cannot give an exact one. Thus, an alternative explanation of the observed contrast, one that does not rely on the assumption that it is unclear whether an exact answer is expected in how many contexts, would be preferable in our view.

Third, contrasts like those in (16) seem difficult to explain on the proposed account:

(16) a. I grew up with more than two parents.
   b. ??I grew up with at least three parents.

Intuitively, (16b) seems odd to us because it has the unlikely implication that the speaker does not remember with how many parents she grew up. On the other hand, (16a) does not seem to force this ignorance implication and is therefore not odd. On W&B’s account, if these sentences are considered out of the blue, an ignorance implicature arises only if the interpreter considers it likely that they were uttered in a context in which a precise answer was required. In this specific case, however, the fact that it is unlikely that the speaker does not remember with how many parents she grew up should make it unlikely for the interpreter that the sentence was uttered in a context in which such an ignorance implicature would have arisen. So the most plausible assumption to make is that the context of utterance did not require an exact answer. But this makes it impossible to explain the oddness of (16b). More generally, we suspect that the ignorance implicature of at least is more robust than predicted by W&B’s theory.

Finally, it should be noted that the polar contexts that W&B considered are quite special, because they involve responses that completely ‘echo’ the question, as exemplified in (17).
This ‘echoing’ may have certain idiosyncratic effects. In order to draw general conclusions about the behaviour of superlative and comparative modifiers in polar contexts it would be preferable to avoid such effects. This may be done by considering dialogues such as (18).

(18) Context: Bill’s diet prescribes that he eat at most three apples per day.
A: Did Bill stick to his diet today?
B: No, he ate more than three apples.
B’: No, he ate at least four apples.

W&B predict that in such dialogues there is no contrast between superlative and comparative modifiers: neither is expected to convey ignorance. M&M, on the other hand, predict that at least does convey ignorance in this case, while more than does not, in line with the received view. This, then, is a crucial case to test in order to assess W&B’s proposal.

3 An approach based on inquisitive sincerity

Büring (2008) expressed the intuition that at least n amounts to a disjunction of the form n or more at some level of analysis. Coppock and Brochhagen (2013b) (C&B) propose that the relevant level of analysis is semantic (rather than syntactic or pragmatic) and show that such an account can be formulated in inquisitive semantics (Ciardelli et al., 2013). In this framework, every sentence generates a set of semantic alternatives, where each semantic alternative is a set of possible worlds. If a sentence generates two or more alternatives, it is thought of as expressing an issue as to which of these alternatives holds. In this case the sentence is called inquisitive. For instance, a disjunctive clause ‘ϕ or ψ’ typically generates a set containing multiple alternatives, namely those associated with ϕ as well as those associated with ψ, and expresses the issue as to which of these holds. According to C&B, the commonality between disjunctions and at least sentences is that both generate multiple semantic alternatives and therefore express non-trivial issues.

More precisely, C&B propose that the set of alternatives generated by an at least sentence consists of all answers to the QUD that are at least as strong as one of the alternatives generated by the prejacent. To illustrate this, first consider the QUD in (19).

(19) How many apples did Bill eat?

Suppose that the answers to this QUD are [0], [1], [2], etcetera. Then we get that:

(20) \[ \text{Bill ate at least four apples} \] = \{[4], [5], [6], \ldots \}

As for more than sentences, C&B assume that they always generate a single alternative, which is the union of the answers to the QUD that are at least as strong as one of the alternatives generated by the prejacent. Thus, in the context of the QUD in (19), we get that:

(21) \[ \text{Bill ate more than three apples} \] = \{[4, \ldots)\}

Turning now from semantics to pragmatics, Coppock and Brochhagen do not only assume the standard Gricean maxims, but also an inquisitive sincerity maxim, which can be characterized informally as: ‘Don’t utter an inquisitive sentence if you already know how to
resolve the issue that it expresses’ (Groenendijk and Roelofsen, 2009). More formally, if a sentence generates multiple alternatives, then the speaker’s information state should not be contained in any of these alternatives. This sincerity maxim, together with the semantics in (20), derives ignorance implicatures for at least sentences in the context of how many questions.

C&B do not explicitly discuss cases where the QUD is a polar question, but it is clear what their account predicts in such cases. Consider the context in (18), repeated in (22).

(22) Context: Bill’s diet prescribes that he eat at most three apples per day.
    QUD: Did Bill stick to his diet today?

Assume that the answers to the QUD are \([0,\ldots,3]\) and \([4,\ldots)\). Then we get that:

(23) \([\text{Bill ate at least four apples}] = \{[4,\ldots]\}\)
(24) \([\text{Bill ate more than three apples}] = \{[4,\ldots]\}\)

That is, both at least and more than just generate a single alternative in this case. Ignorance implicatures are therefore not predicted to arise, at least not through the inquisitive sincerity maxim. C&B do not explicitly discuss whether ignorance implicatures may arise through Gricean quantity-based reasoning. But as far as inquisitive sincerity-based ignorance implicatures are concerned, they predict that more than has none, and that at least has them in how many contexts but not polar contexts.

This concludes our discussion of some of the existing approaches to superlative and comparative modified numerals. As anticipated in the introduction, we have seen that these approaches differ both in empirical predictions and in what they take to be the fundamental source of ignorance implicatures triggered by modified numerals. In the next section, we provide novel data from a series of experiments, aiming to establish a more solid empirical basis for evaluating these approaches and further developing them.

4 Experimental data

4.1 Experiment 1

4.1.1 Goals
The goal of our first experiment was to establish some basic facts regarding ignorance implicatures triggered by more than and at least in the context of two types of QUD: how many questions and polar questions. In particular, we wanted to see whether the patterns found by Westera and Brasoveanu (2014), reviewed above, could be replicated using a different experimental setup, one that is closer to the truth-value judgment task standardly used in the experimental literature on implicatures.

Participants had to judge the acceptability of statements made during a card game. This setup made it very easy to construct complete information and ignorance scenarios. In the former kind of scenario, a player makes a statement about her cards at a point in the game where she knows what all her cards are. In the latter, she makes a statement about her cards at a point where she has not seen all her cards yet. Both scenarios can be represented visually, by means of a simple picture (see Figure 3 for an example).

In this setup, we can test the presence of ignorance inferences by studying how people judge interpretations that violate such inferences. This is a common strategy in other
pragmatic studies, for example, on scalar implicatures (Bott and Noveck, 2004). The specific design that we used has been employed successfully in a study of primary scalar implicatures in Dieuleveut et al. (2017).

Given that the scenario is made very explicit, it is possible to ensure that polar questions and how many questions have exactly the properties that we want them to have. In particular, we can specify details of the scenario in such a way that it is extremely unlikely that how many questions would be treated underlingly as polar questions.

We can also avoid echoic responses to polar questions that W&B used, capitalizing on the fact that I have at least six clubs entails a ‘yes’ answer to the question Will you win this round? (see below for details). This is desirable because echoic responses may lack the semantic alternatives that they would normally have.

Finally, our setup allows us to test differences between answers which exactly match the ‘yes’ answer to a polar question and answers which provide more information than necessary.

4.1.2 Methodology

Participants: 50 participants were recruited on MTurk for the experiment. The participants were self-reported adult English speakers (age: 18 – 67, mean: 35). They received $2.50 for their participation in the experiment.

Materials and procedure: The task of every participant was to judge the appropriateness of statements on a 5-point scale (1 signalled the lowest level of appropriateness, and 5 signalled the highest level). Statements were presented as answers to questions and were accompanied by a pictorial situation. An example of an item is given in Figure 3.

The following background story preceded the actual experiment:

Mary is playing a card game online. Each round consists of a betting phase and a playing phase.
**Betting phase:** At the beginning of the betting phase, each player receives 8 cards: 6 visible (to them) and 2 face down. Players look at their first six cards. Based on what they see, all players place a bet on their chance to win the round. Then the players see their seventh card and place a second bet. A final bet is placed after players see their eighth and last card. When all players have placed their last bets, the playing phase begins.

**Playing phase:** The exact rules for this part of the game are not relevant here. Clubs are the trump suit (a club card beats any card from other suits), and if Mary receives six or more clubs, she is bound to win (she has an unbeatable strategy).

**Scoring:** If Mary wins the round, she earns the number of points she bet, plus one bonus point for each face card she had (Jack, Queen or King). Face cards only influence scoring and do not play any special role in the playing phase.

Sue is Mary’s best friend. She knows the rules of the game and Mary’s strategy, but is not currently playing. Sue walks into the room at some point during the betting phase and asks Mary a question. Sue cannot see Mary’s screen, so she doesn’t know which cards Mary got or whether Mary has seen all of her cards yet when she asks her question. Mary has no reason to hide information from Sue.

Each stimulus consisted of the picture of 8 cards, showing Mary’s cards. Sue’s question and Mary’s answer were below the picture (see Fig. 3).

The experiment consisted of 108 experimental stimuli and 51 fillers. The experimental stimuli tested 3 quantifiers (Condition: QUANTIFIER), 3 QUD-Answer relations (Condition: QUD) and 4 situations (Condition: SITUATION). Each combination of conditions was repeated three times.

The following three quantifiers were used for the QUANTIFIER condition: *at least* $n$, *more than* $n$, and the bare numeral $n$. For example, for the situation depicted in Fig. 3, these three versions were used:

\[(25)\]

a. At least three of my eight cards are clubs.
b. More than two of my eight cards are clubs.
c. Three of my eight cards are clubs.

The QUD manipulation consisted of 3 conditions. In the first (HOW MANY), Sue asked a how many question about the number of face cards (see Fig. 3). The other two conditions involved polar questions, and differed with respect to whether the answer was directly relevant to the question. In the POLARRELEVANT condition, Sue asked a polar question (*Will you win this round?*) and Mary responded using the relevant information to determine winning (e.g., Yes, at least six/more than five of my eight cards are clubs). In the POLAROVERINF condition, Sue asked the same polar question and Mary responded providing more information than necessary (e.g., Yes, at least seven/more than six of my eight cards are clubs).

The SITUATION manipulation consisted of 4 conditions. In the first case (FALSE), all the cards were revealed and Mary’s answer was simply false: there were always 2 clubs/face cards less than what she reported. In the second case (IGNORANCE), Mary’s answer was given during the uncovering phase (one or two cards were still covered). The cards revealed

---

3Recall that in order to be sure she will win, Mary had to have 6 or more clubs.
so far supported the lower bound of her response. In the third case (EXCEED), all the cards were revealed and there was one more club/face card than the lower bound conveyed by Mary’s answer. In the last condition (EXACT), all the cards were revealed and the number of club/face cards was exactly the lower bound conveyed by the answer.

The actual experiment was preceded by 6 practice trials. The practice trials were similar to the experimental stimuli in their design, but unlike experimental items they indicated whether Mary’s answer was appropriate or not and why this was so. None of the practice trials used any construction that was tested in the experiment (at least or more than).

**Pre-processing and data analysis:** We first removed the fastest and slowest 1% of responses as outliers. We then calculated individual error rates on control items for which we expected a clear answer. In such cases, ‘3’ never counted as a correct answer (hence the theoretical chance level is at 40%). Four participants were removed because their error rate on control items was at least one standard deviation above the mean error rate (threshold: 22.7%). The mean error rate on remaining participants was 7.6%.

For statistical analysis, responses were treated as a continuous variable and normalized by participant. We fitted mixed-effects linear models with the lme4 package in R (R Core Team, 2014; Bates et al., 2014), following the recommendations of Bates et al. (2015) regarding the specification of the random effects structure. For the calculation of p-values, we approximated the t-distribution with a Gaussian curve. (This approximation should not be problematic given the number of participants.)

### 4.1.3 Results

The detailed results on target conditions are presented in Figure 4. In the rest of this section and for the statistical analyses, we leave aside the FALSE situations, which gave rise to very low ratings in all conditions, as expected.

We fitted a linear mixed-effects model on responses to the EXACT, EXCEED and IGNORANCE conditions for sentences with at least and more than, in response to how many and polar questions. All three factors (QUANTIFIER, QUD, and SITUATION) were treatment-coded, with IGNORANCE as the SITUATION baseline, more than as the QUANTIFIER baseline, and how many as the QUD baseline. The results are given in Table 3.

The middle column in Figure 4 shows that, across all QUD conditions (corresponding to the rows), ratings for more than sentences in the EXACT, EXCEED, and IGNORANCE situations are all roughly on a par with each other. Thus more than does not appear to exhibit a preference for ignorance situations. This conclusion is supported by our statistical analysis shown in Table 3. With more than sentences, across all QUD conditions, the differences between the baseline IGNORANCE condition and the non-IGNORANCE conditions are non-significant (with one possible exception: for sentences giving a relevant answer to a polar QUD, the ratings were 0.41 scale points higher on average in the EXCEED situation, compared to the IGNORANCE situation, and this difference is significant, though only at the 0.05 level, which is arguably too lax given the number of comparisons we are testing).

Things were different in the case of at least, as seen in the leftmost column in Figure 4, showing that ratings for at least sentences in non-IGNORANCE situations were generally lower than in IGNORANCE situations. This is supported by the statistical analysis. Across all QUD types, the ratings for sentences in the EXACT condition were significantly lower compared to the IGNORANCE baseline (between 0.82 and 1.13 scale points lower on average). The finding supports previous claims that precise knowledge negatively affects the acceptability of at least more than the acceptability of more than.
The relative acceptability of at least in the EXCEED situation, compared to the IGNORANCE situation, depended on the nature of the QUD. Ratings for at least sentences were significantly lower in the EXCEED situation in the context of a how many question (0.66 scale points lower on average), but not in the context of a polar question (0.29 or 0.32 scale points lower on average, depending on whether the answer was relevant). Hence, broadly speaking, the deviance due to ignorance inferences of at least is not as strong in the context of polar questions.

As a post-hoc analysis, we fitted a model on the results of at least which pooled the EXACT and EXCEED situations together as a new condition PRECISE. The POLARRELEVANT and POLAROVERINF conditions were also pooled together as a new condition, POLAR. The model did not explain significantly less variance than a full model ($\chi^2(5) = 8.5, p = 0.13$), and it showed a strong negative effect of the PRECISE situation ($t = -5.2, p < .001$) as well as a strong positive interaction of the PRECISE situation with the POLAR QUD ($t = 4.9, p < .001$).

We analyzed the results for bare numerals separately. Unsurprisingly, we observed that the EXACT condition was rated higher than the IGNORANCE condition ($t = 5.2, p < .001$), which in turn was higher than the EXCEED condition ($t = 13, p < .001$). The FALSE condition was still lower than the EXCEED condition ($t = 5.7, p < .001$). Interestingly, the EXCEED and IGNORANCE conditions were both rated higher with POLAR QUDs ($t = 3.7$ and $t = 7.8$, respectively, both $p < .001$), while the FALSE and EXACT conditions were unaffected by the QUD type (both $t < .4, p > .72$).

### 4.1.4 Discussion

We saw no trace of ignorance inferences with more than, except for a small trend in responses to how many questions (in the EXCEED situation). By contrast, at least gave rise to clear ignorance inferences, as evidenced by its lower acceptability in EXCEED and EXACT.
Table 3: Estimates for the fixed effects of the model fitted on target conditions for Experiment 1. (AL = at least)

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.05</td>
<td>-0.6</td>
<td>0.58</td>
</tr>
<tr>
<td>Situation[Exact]</td>
<td>-0.07</td>
<td>-0.6</td>
<td>0.57</td>
</tr>
<tr>
<td>Situation[Exceed]</td>
<td>-0.24</td>
<td>-1.8</td>
<td>0.073</td>
</tr>
<tr>
<td>Quantifier[AL]</td>
<td>0.42</td>
<td>3.2</td>
<td>&lt; 0.001**</td>
</tr>
<tr>
<td>QUD[PolarRelevant]</td>
<td>-0.11</td>
<td>-0.8</td>
<td>0.44</td>
</tr>
<tr>
<td>QUD[PolarOverInf]</td>
<td>0.41</td>
<td>3.5</td>
<td>&lt; 0.001**</td>
</tr>
<tr>
<td>Situation[Exact]:Quantifier[AL]</td>
<td>-1.13</td>
<td>-5.5</td>
<td>&lt; 0.000***</td>
</tr>
<tr>
<td>Situation[Exceed]:Quantifier[AL]</td>
<td>-0.66</td>
<td>-3.8</td>
<td>&lt; 0.000***</td>
</tr>
<tr>
<td>Situation[Exact]:QUD[PolarRelevant]</td>
<td>-0.02</td>
<td>-0.1</td>
<td>0.93</td>
</tr>
<tr>
<td>Situation[Exceed]:QUD[PolarRelevant]</td>
<td>0.41</td>
<td>2.3</td>
<td>0.019*</td>
</tr>
<tr>
<td>Situation[Exact]:QUD[PolarOverInf]</td>
<td>0.06</td>
<td>0.4</td>
<td>0.69</td>
</tr>
<tr>
<td>Situation[Exceed]:QUD[PolarOverInf]</td>
<td>0.24</td>
<td>1.6</td>
<td>0.11</td>
</tr>
<tr>
<td>Quantifier[AL]:QUD[PolarRelevant]</td>
<td>0.10</td>
<td>0.6</td>
<td>0.54</td>
</tr>
<tr>
<td>Quantifier[AL]:QUD[PolarOverInf]</td>
<td>-0.48</td>
<td>-3.1</td>
<td>&lt; 0.002**</td>
</tr>
<tr>
<td>Situation[Exact]:Quantifier[AL]:QUD[PolarRelevant]</td>
<td>0.90</td>
<td>3.8</td>
<td>&lt; 0.000***</td>
</tr>
<tr>
<td>Situation[Exceed]:Quantifier[AL]:QUD[PolarRelevant]</td>
<td>0.29</td>
<td>1.3</td>
<td>0.20</td>
</tr>
<tr>
<td>Situation[Exact]:Quantifier[AL]:QUD[PolarOverInf]</td>
<td>0.82</td>
<td>3.5</td>
<td>&lt; 0.000***</td>
</tr>
<tr>
<td>Situation[Exceed]:Quantifier[AL]:QUD[PolarOverInf]</td>
<td>0.32</td>
<td>1.6</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Note that the absence of ignorance inferences with more than cannot be explained by a lack of statistical power, as we were able to clearly detect QUD effects on at least. If anything, the standard errors for our data seem to be overall smaller than W&B’s. Our experimental design differed from theirs on a few key points, which could explain the observed differences. First of all, our participants’ task was to evaluate the appropriateness of an utterance given the knowledge state of a speaker, while W&B required participants to judge an inference about a speaker’s knowledge state, given their utterance. In short, we could say that our task was more “speaker oriented” (evaluating whether an utterance is acceptable is similar to considering whether one could utter such an utterance) while W&B’s task was more “hearer oriented”. Second, we situated our conversations in a more casual set-up (a discussion between friends), while W&B used a very formal context (a witness testifying in front of a judge). Third, we tested positive modified numerals (at least and more than), whereas W&B tested their negative counterparts (at most and less than). Coppock and Brochhagen (2013a) tested both positive and negative modified numerals and found a difference between them: truth-value judgments for at most n were strongly degraded for situations depicted with fewer than n objects, in contrast to fewer than n + 1; at least n + 1 and more than n did not show such a strong contrast. If anything, the difference between the results of our first experiment and W&B go in the opposite direction. The influence of monotonicity is thus unclear.

At this point, it is unclear which of these factors may explain the discrepancies between our results and those reported by W&B. In the next two experiments, we adopt a design much closer to theirs, while avoiding the issues we identified in their experiments. To foreshadow the results, the first factor (“speaker” vs. “hearer” orientation) seems to best
explain the differences, but a contrast between positive and negative modified numerals may have played a role as well.

Before turning to the next experiment, let us briefly discuss the results obtained with bare numerals, even though this is not the main focus of the paper. In general, our findings support the view that numerals start with a one-sided (at-least) denotation and receive their exact reading through implicatures (Horn, 1972, see Spector, 2013 for a review). First, we see that the Exceed condition (which involved a situation with more than \( n \) items) is rated much higher than the False condition (which involved a situation with less than \( n \) items), suggesting that bare numerals are more sensitive to violations in one direction than in the other. Second, the fact that Exceed is less accepted than Ignorance (where the exact reading is neither supported nor excluded) is indicative of a primary implicature, as predicted by the traditional implicature approach. The fact that Ignorance is less accepted than Exact could be indicative of either secondary implicatures or an ambiguity between one-sided and two-sided readings. Third, we found that the contrasts between Exceed and Ignorance and between Ignorance and Exact were sensitive to the QUD. This is at odds with the ambiguity approach of Geurts (2006) or Kennedy (2015), which assumes that the ambiguity is always resolved in favor of the strongest reading (Strongest Meaning Hypothesis). Overall, these results are surprising because in most experimental settings, numerals diverge from archetypal scalar implicatures (see, e.g., Papafragou and Musolino, 2003 for acquisition, Huang and Snedeker, 2009 for processing). In particular, Dieuleveut et al. (2017), using a task very similar to ours, found primary implicatures with the quantifier some but not with numerals.\(^4\)

### 4.2 Experiment 2

#### 4.2.1 Goals

The main goals of our second experiment were (a) to understand why the results of our first experiment differed so sharply from W&B’s results, and (b) to attempt once more, in a different experimental setting, to detect ignorance inferences with comparatives, in order to test their possible QUD-dependency. For this purpose, we adopted a design very close to W&B’s. In particular, we switched to an inferential task, in which participants had to evaluate how much a speaker knew given her answer to a question.

The characters involved in this experiment were police officers or investigators (who asked questions) and witnesses (who responded to these questions). This brings us closer to W&B’s judge/witness situation, while allowing more variety in the situations we considered. The conversation was also likely to be more casual than in a courtroom.

As discussed above, one other possible source of the discrepancies between the results of our first experiment and those of W&B, is the fact that we considered positive modifiers (at least and more than) while they considered negative ones (at most and less than). In our second experiment, we tested both positive and negative modifiers.

Finally, as in our first experiment, we wanted to get a better understanding of the potential contrast between propositions which match a complete answer to a given polar question and propositions which are over-informative. In the previous experiment there was a possible confound: we saw that the acceptability of more than was overall degraded when it combined with a numeral which was not salient, even though it provided a complete answer to the question, whereas the over-informative condition was overall well accepted,

\(^4\)Something that might explain this difference is that Dieuleveut et al. (2017) used a binary response option while we collected graded judgments.
possibly because it involved the salient numeral *six*. In our second experiment, we varied the polar question in such a way that *more than n* was sometimes but not always over-informative, while keeping *n* salient. How this was achieved is explained in more detail below.

### 4.2.2 Manipulating relevance through context

We manipulated the relevance of the response provided by a modified numeral to a polar question by varying contexts instead of the numeral (as we did in Experiment 1). Schematically, if *n* is the numeral that was modified, “upward” contexts involved a polar question equivalent in terms of resolution conditions to “is it at least *n*?”, while “downward” contexts involved a polar question equivalent to “is it at most *n*?”.

However, we needed to make sure that the questions would not explicitly contain a modified numeral. Upward contexts typically established a requirement for *n* items, whereas downward contexts specified a maximum number of items. This way, we could simply ask whether the rules had been respected, without having to mention any modified numeral in the question.

Let us illustrate this with some concrete examples. In an upward context, an investigator may ask whether there were enough seat belts for the five passengers in a car. If the witness answers that there were *more than five* seat belts, she gave an over-informative positive answer. *Fewer than five* would be a negative relevant answer. Finally, *at least five* would be a positive relevant answer, but *at most five* would not resolve the question (there may have been five, and there may have been fewer than five).

The roles are reversed in a downward context, in which an investigator may for instance ask whether the maximum load of ten people was exceeded during an elevator incident. In this case, saying that *more than ten* people were present in the elevator is a positive relevant answer, while *fewer than ten* is a negative over-informative answer. Finally, *at most ten* is a negative relevant answer, while *at least ten* does not resolve the question.

To sum up, our manipulation allowed us to test comparatives both as relevant and as over-informative answers, and the numeral being modified was always salient and round, whether the resulting construction matched a complete answer or not. Superlatives always corresponded to relevant answers (cases in which superlatives did not answer the question were excluded from the experiment, since they would introduce an orthogonal issue).

### 4.2.3 Methodology

**Participants:** 95 participants were recruited on MTurk for the experiment. The participants were self-reported adult English speakers (age: 20 – 63, mean: 35). They received $1.80 for their participation in the experiment.

**Materials and procedure:** Every participant read the following instructions on the first screen:

> In this survey, you will see short dialogues between police officers and witnesses in some legal cases. Each example will come with a few sentences giving the context of this discussion. The witnesses are neither suspects nor plaintiffs, so they have a neutral position in the cases. They have no reason to hide

---

5Concretely, the questions we used were phrased positively, so that positive quantifiers would always be associated with the response particle ‘Yes’, while negative ones were associated with the particle ‘No’. In this sense, they are closer to “is it more than *n*?” than “is it at most *n*?”.
information from the investigators, and are therefore being as cooperative as they can.

Each experimental item consisted of a context story, a question-answer pair, and a prompt. An example is presented in Figure 5. We manipulated the following factors: QUD (polar or how-many), Quantifier (superlative, comparative-relevant, comparative-over-informative, and bare numeral), and CONTEXT_TYPE (upward or downward). All factors were within subject. All factors but CONTEXT_TYPE were within item.

The QUD factor determined whether the investigator’s question was a polar or a how-many question. Each context story came in two versions: one that mentioned an explicit threshold, for the polar QUD, and one which did not specify any threshold, for the how-many QUD. Responses to polar questions always involved the response particle (‘yes’ or ‘no’ depending on the case). The concrete quantifier used in the witness’s response depended on both Quantifier and CONTEXT_TYPE, as explained above: in upward contexts, the superlative modifier was at least, the relevant comparative was fewer than, and the over-informative comparative was more than. In downward contexts, the superlative modifier was at most and the roles of fewer than and more than were reversed. Note that this contrast between over-informative and relevant answers only makes sense for polar questions, but to keep the comparison minimal, we kept the same quantifiers in how-many QUDs. The context fixed the numeral used in all conditions. It was always a round number, ranging from 5 to 1000, and it was written in words.

In all target items, the prompt was a question of the form “Would you conclude that the witness knows exactly how many . . . ?”. Participants responded using a 5-point Likert scale from “Definitely not” to “Definitely yes”.

We designed 24 contexts (12 upward, 12 downward). The 8 possible combinations of QUD and Quantifier were presented three times to each participant following a latin-square design. This means that some participants would see some combinations of quantifiers and QUD as upward twice and downward once, while other participants would see the same combinations as downward twice and upward once. The order was fully randomized across participants.

In addition to the 24 targets, each participant saw the same 10 fillers (5 true, 5 false).

Pre-processing and data analysis: Pre-processing was done in the same way as for Experiment 1. We decided to ignore three fillers when computing error rates because they tested participants’ attention in a convoluted way and were answered incorrectly in 90% of cases on average. Eleven participants were removed because their error rate on the rest
Figure 6: Individual responses to each target condition in Experiment 2. The relevance of the answer provided by a comparative modified numeral to the polar question, marked by the blue-green contrast here, depended on an interaction between context type (upward or downward) and polarity (negative fewer than or positive more than).

of filler items was at least one standard deviation above the mean error rate (threshold: 31.8%). The mean error rate on remaining participants was 6.1%. The analyses followed the methods of the first experiment, with the only difference that the used mixed-effects models also had items random effects (in addition to subjects random effects).

4.2.4 Results

The results are presented in Figure 6. The higher the response, the less ignorance participants attributed to the witness.

For the statistical analysis, we defined a Polarity factor: at least and more than are positive quantifiers, whereas at most and fewer than are negative quantifiers. We first ran a model on responses to all target items with the following predictors: QuantifierType (comparative vs. superlative), Polarity (centered, with negative at −0.5, and positive at +0.5) and QUD (centered, with how-many at −0.5, polar at +0.5). The results, given in Table 4, showed significant main effects of Polarity (positive quantifiers give rise to less ignorance) and QUD (polar QUDs give rise to less ignorance). These two effects interacted positively, suggesting that the effect of QUD was stronger for positive quantifiers. No other effect was significant, in particular, none of the interactions associated with QuantifierType.

We fitted a second model focused on data from comparative modifiers and polar QUD to test for effects of relevance. This model confirmed the results of the first model (clear main effects of Polarity), but showed no effect of Relevance (t = −0.21, p = .83) and no interaction (t = 1.5, p = .13: if anything more than gives rise to less ignorance when providing an over-informative answer, which is the opposite of what one would expect).

We also fitted a model on the results involving bare numerals. This showed a marginal effect of ContextType (t = 1.96, p = .05), but no significant effect of QUD or interaction (both |t| < 1.1, p > .30).
Table 4: Estimates for the fixed effects of the model fitted on target conditions for Experiment 2.

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.59</td>
<td>-16.72</td>
<td>0.000***</td>
</tr>
<tr>
<td>QuantifierType[Superlative]</td>
<td>0.06</td>
<td>1.30</td>
<td>0.20</td>
</tr>
<tr>
<td>Polarity</td>
<td>0.24</td>
<td>3.84</td>
<td>0.000***</td>
</tr>
<tr>
<td>QUD</td>
<td>0.18</td>
<td>4.01</td>
<td>0.000***</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Polarity</td>
<td>0.12</td>
<td>1.41</td>
<td>0.16</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:QUD</td>
<td>0.05</td>
<td>0.70</td>
<td>0.48</td>
</tr>
<tr>
<td>Polarity:QUD</td>
<td>0.24</td>
<td>2.69</td>
<td>0.007**</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Polarity:QUD</td>
<td>-0.04</td>
<td>-0.29</td>
<td>0.77</td>
</tr>
</tbody>
</table>

4.2.5 Discussion

We replicated Coppock and Brochhagen's (2013a) finding that negative quantifiers give rise to stronger ignorance inferences than positive ones. We did not observe any effect of relevance on responses to polar QUDs. Recall that in Experiment 1, we observed that for more than, POLAROVERINF was judged differently than POLARRELEVANT. This, we conjectured, could be due to the difference in relevance, or due to the fact that POLAROVERINF required mentioning a number that was salient in that context (more than six), unlike POLARRELEVANT. The null effect of relevance in this experiment suggests that the contrast observed in Experiment 1 was caused by the salience of six.

We also confirmed that ignorance inferences are QUD-sensitive, as we had observed in Experiment 1. However, unlike in Experiment 1, we observed no difference between superlative and comparative modified numerals regarding the strength of ignorance inferences. This is in line with the results of W&B, and it raises an important question: what could explain the difference between our first experiment on the one hand, and W&B and our second experiment on the other hand? The third experiment addressed this issue.

4.3 Experiment 3

4.3.1 Goals

The main goal of this experiment was to gain a better understanding of the apparent conflict between the results of the first two experiments. We hypothesized that, among the various factors we identified in Section 4.1.4, the difference in the task was the most likely explanation. In Experiment 1, participants had to judge the acceptability of an utterance, while they were informed about the knowledge state of the speaker. In Experiment 2, participants had to decide whether or not to draw an ignorance inference, based on the context and the given utterance. In order to directly assess the hypothesis that the difference in task was responsible for the observed contrasts between the results of the two experiments, we adapted the task of Experiment 2, making it very similar to the one of Experiment 1. At the same time, the materials of Experiment 2 were preserved as much as possible. How this was done is described in more detail below.

4.3.2 Methods

Participants: 96 participants were recruited on MTurk for the experiment. The participants were self-reported adult English speakers (age: 20 – 67, mean: 35). They received $1.50 for their participation in the experiment.
Figure 7: Example of a target Approximate item in Experiment 3 with a downward context, polar QUD, and more than.

<table>
<thead>
<tr>
<th></th>
<th>‘at least n’</th>
<th>‘more than n’</th>
<th>‘at most n’</th>
<th>‘fewer than n’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precise:</td>
<td>n + δ</td>
<td>between n + δ</td>
<td>n - δ</td>
<td>between n - 3δ</td>
</tr>
<tr>
<td>Approximate:</td>
<td>between n</td>
<td>between n + δ</td>
<td>n - 3δ</td>
<td>between n - 3δ</td>
</tr>
<tr>
<td></td>
<td>and n + 3δ</td>
<td>and n + 3δ</td>
<td>and n</td>
<td>and n - δ</td>
</tr>
</tbody>
</table>

Table 5: Description of the witness’s knowledge in each condition for each quantifier, where n is the numeral used by the witness and δ is a lower level of granularity (e.g., n = 10 and δ = 1, or n = 500 and δ = 25).

Materials and procedure: The experiment mainly differed from Experiment 2 in the following two respects. First, we added information to each item on what the witness actually knew. Second, we changed the prompt from “Would you conclude that the witness knows...” to “Is the witness’s answer appropriate?”. An example is given in Figure 7.

We added a two-level factor Knowledge determining how much the witness knew about each situation. In the Precise condition, the witness knew exactly what the number of relevant items was. This would be in conflict with potential ignorance inferences triggered by her answer. In the Approximate condition, the witness only knew a range of possible values that was chosen to be maximally compatible with a potential ignorance inference. The numbers were selected based on the algorithm shown in Table 5.

The addition of the Knowledge manipulation would have doubled the number of items in comparison with Experiment 2. We dropped certain conditions to compensate for this: since we did not observe any difference between relevant and over-informative answers in Experiment 2, we kept only relevant answers this time. This means that in the polar QUD cases, more than was not tested in upward contexts, and fewer than was not tested in downward contexts.

The modifiers and the QUDs were balanced across contexts, although their combinations were not: at least and fewer than appeared in the HowMany condition of downward contexts, and at most and more than appeared in the HowMany condition of upward contexts. This was because, as explained in the method section of the previous experiment (Section 4.2.2), in case of polar QUDs not every context could appear with every modifier. That is, only upward contexts could be combined with at least and fewer than and downward contexts with at most and more than. Since how many questions do not pose such a restriction, they were used to balance the design. This move was justified by a post-hoc

---

6In addition to being compatible with potential ignorance inferences, the range of possible values in the Approximate condition respected the exhaustivity inference discussed in Cummins et al. (2012) as well.
analysis of data from the previous experiment, showing that CONTEXTTYPE did not have any interaction with QUD or POLARITY ($\chi^2(3) = .88, p = .83$). Bare numerals appeared with both QUDs in each context.

We used a latin-square design with 24 context stories (the same stories as in Experiment 2). The experiment also included 2 training items (immediately after the instructions) and 12 fillers (8 were false controls, where the witness’s knowledge directly contradicted her statement; in the other 4 the witness had approximate knowledge and used the expression about $n$ in her response).

**Pre-processing and data analysis:** Pre-processing and analysis were done in the same way as for Experiment 2. Ten participants were removed because their error rate on the filler items was at least one standard deviation above the mean error rate (threshold: 30.8%). The mean error rate on remaining participants was 2.7%.

4.3.3 Results

The results are presented in Figure 8. We ran a model on responses to all target items with POLARITY, QUD, and QUANTIFIERTYPE encoded as in Experiment 2: QUANTIFIERTYPE was treatment-coded (comparative vs. superlative), POLARITY was centered (with negative at $-0.5$, and positive at $+0.5$) and QUD was centered (with how-many at $-0.5$, and polar at $+0.5$). KNOWLEDGE was treatment-coded with Approximate as the baseline. The results, given in Table 6, showed a significant effect of QUANTIFIERTYPE (superlative quantifiers are more acceptable in Approximate situations than comparatives are), an effect of QUD (the acceptability of comparatives is higher with polar QUDs), a strong interaction between KNOWLEDGE and QUANTIFIERTYPE (superlative quantifiers are clearly degraded in Precise situations), and a triple interaction between KNOWLEDGE, QUANTIFIERTYPE and POLARITY (at most is more sensitive than at least to ignorance violations). Most notably, KNOWLEDGE had no effect whatsoever on comparative quantifiers, showing that these quantifiers do not convey ignorance in this experimental setup.

<table>
<thead>
<tr>
<th></th>
<th>$\beta$</th>
<th>$t$-value</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.56</td>
<td>13.50</td>
<td>0.000***</td>
</tr>
<tr>
<td>QuantifierType[Superlative]</td>
<td>0.15</td>
<td>2.82</td>
<td>0.005**</td>
</tr>
<tr>
<td>Polarity</td>
<td>-0.04</td>
<td>-0.42</td>
<td>0.674</td>
</tr>
<tr>
<td>Knowledge[Precise]</td>
<td>0.02</td>
<td>0.45</td>
<td>0.654</td>
</tr>
<tr>
<td>QUD</td>
<td>0.27</td>
<td>2.46</td>
<td>0.014*</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Polarity</td>
<td>0.14</td>
<td>1.31</td>
<td>0.191</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Knowledge[Precise]</td>
<td>-0.59</td>
<td>-7.90</td>
<td>0.000***</td>
</tr>
<tr>
<td>Positive:Knowledge[Precise]</td>
<td>0.04</td>
<td>0.30</td>
<td>0.762</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:QUD</td>
<td>-0.14</td>
<td>-0.96</td>
<td>0.335</td>
</tr>
<tr>
<td>Polarity:QUD</td>
<td>0.18</td>
<td>1.27</td>
<td>0.203</td>
</tr>
<tr>
<td>Knowledge[Precise]:QUD</td>
<td>-0.02</td>
<td>-0.15</td>
<td>0.877</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Polarity:Knowledge[Precise]</td>
<td>0.65</td>
<td>4.64</td>
<td>0.000***</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Polarity:QUD</td>
<td>-0.03</td>
<td>-0.16</td>
<td>0.876</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Knowledge[Precise]:QUD</td>
<td>0.27</td>
<td>1.57</td>
<td>0.117</td>
</tr>
<tr>
<td>Polarity:Knowledge[Precise]:QUD</td>
<td>-0.15</td>
<td>-0.80</td>
<td>0.426</td>
</tr>
<tr>
<td>QuantifierType[Superlative]:Polarity:Knowledge[Precise]:QUD</td>
<td>-0.08</td>
<td>-0.30</td>
<td>0.765</td>
</tr>
</tbody>
</table>

Table 6: Estimates for the fixed effects of the model fitted on target conditions for Experiment 3.
Figure 8: Individual responses to each target condition in Experiment 3. Similar to Experiment 1, ignorance inferences manifest themselves as differences between the Precise and Approximate conditions.

Overall, these results are very close to what we found in Experiment 1: strong ignorance effects with superlative modifiers, and none with comparative modifiers. Interestingly, the three-way interaction between Knowledge, QuantifierType and QUD, which would have indicated QUD sensitivity for the ignorance inferences of superlative modifiers, did not reach significance ($p = .11$). While this might seem surprising at first sight, this finding also matches the results of Experiment 1. There, we observed a three-way interaction, indicating that the ignorance inferences of superlative modifiers are QUD-sensitive. However, this three-way interaction was only present in the Exact Situation, while in the Exceed Situation, it was of the same magnitude as in the current experiment and failed to reach significance. Since the Precise Knowledge condition in this Experiment is closer to the Exceed Situation in Experiment 1, this result is not surprising.

In line with Experiment 2, ignorance inferences affected the acceptability of the negative modifier at most more than that of the positive modifier at least.

4.3.4 Discussion

Figure 9 presents a condensed summary of the most important results obtained in the three experiments. The graphs highlight the main finding: in Experiments 1 and 3, the acceptability of superlative modifiers was sensitive to ignorance effects, unlike that of comparative modifiers. In Experiment 2, we observed no difference between the two types of modifiers. In all three experiments, we observed that whenever an effect of ignorance was present, it was sensitive to the QUD. Furthermore, negative quantifiers systematically gave rise to stronger ignorance inferences than their positive counterparts.

The results of Experiment 3 provide clear indications on how to reconcile the conflicting results obtained in the previous two experiments. Diverging minimally from the setup in
Experiment 2, we were able to retrieve results very similar to those of Experiment 1. We thus conclude that the source of the contrast observed between the results of Experiments 1 and 2, respectively, indeed lies in the difference between the two tasks: (i) when participants judge the appropriateness of an expression while having access to the speaker’s knowledge, they only take superlative modifiers to signal ignorance; (ii) when they have to draw inferences about the speaker’s knowledge given what the speaker said in a certain context, they assume both superlative and comparative modifiers to signal ignorance, which is modulated by QUDs.

Here is one way to interpret this finding. The two tasks make participants take different perspectives in communication. In Experiment 1 and 3, they are more likely to take the speaker’s perspective, as they have access to her internal mental state and judge the appropriateness of different utterances. This makes the task relatively close to actual production. In Experiment 2 and W&B’s experiment, participants are more likely to take the hearer’s perspective, as they have to draw inferences about the speaker’s state based on what was said. Our results suggest that when taking the perspective of speakers in the task, participants are sensitive to the comparative/superlative distinction with respect to ignorance inferences, but this distinction does not play a role when they take the hearer’s perspective. A possible interpretation is that participants are sensitive to different sets of maxims in the two tasks, or that they weigh the maxims differently. We will propose in Section 5 that, while the ignorance inference of more than arises through the maxim of quantity, that of at least arises both through quantity and through the independent maxim of inquisitive sincerity. Our findings can then be accounted for under the assumption that when taking the speaker’s perspective, participants give more importance to inquisitive sincerity, while when taking the hearer’s perspective, they give more importance to quantity.\footnote{Lauer (2014) argues that violations of quantity give rise to “Need a reason” implicatures, which often – but not always – amount to inferring that the speaker is ignorant. That is, when hearing an utterance which had a more informative alternative, we look for a reason why the speaker did not utter this alternative, and the most salient candidate explanation is that doing so would have violated Quality (i.e., the speaker}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9.png}
\caption{Mean and SE for a derived measure of Ignorance in each of the 3 experiments. For Experiments 1 and 3, we computed the mean difference between conditions that respected and violated ignorance. For Experiment 2, we computed 3 minus the mean answer. 0 corresponds to a no-ignorance baseline, although it is somewhat arbitrary in the case of Experiment 2. Relevant and over-informative answers to polar questions were pooled together.}
\end{figure}
Apart from their theoretical significance, our findings are also of interest from a methodological point of view: we showed that two different tasks tapped into different aspects of the semantics and pragmatics of modified numerals. If we are right that the differences are caused by the fact that participants weigh maxims differently when considering their responses, this would have implications for the experimental investigation of implicatures more broadly.

We also observed a systematic contrast between positive and negative quantifiers in Experiments 2 and 3. A similar contrast had been observed in Coppock and Brochhagen (2013a) but has not been explained, as far as we know. A tentative explanation would be that this contrast results from world knowledge on how information is acquired. Concretely, one can gather positive evidence that there are at least five children in a room by counting the visible children and keep open the possibility that some of them may be hiding. By contrast, it is much harder to construct a scenario in which one obtains direct evidence that there are at most five children in a room. Since it is harder to collect direct evidence supporting a downward entailing statement, participants may be tempted to attribute stronger ignorance to speakers who utter such statements.

5 A dual-route approach

We now spell out an account that combines a number of insights from the quantity-based and inquisitive sincerity-based approaches.\(^8\) We provide the necessary background notions from inquisitive semantics in Section 5.1, spell out the semantic component of our account in Section 5.2, and then turn to the pragmatic component in Section 5.3.

5.1 Background notions and notation

In inquisitive semantics the meaning of a sentence \(\varphi\), denoted \(\llbracket \varphi \rrbracket\), is a set of propositions encoding both the information that is conveyed and the issue that is expressed by \(\varphi\). Specifically, \(\varphi\) is taken to convey the information that the actual world is contained in \(\bigcup \llbracket \varphi \rrbracket\), which is denoted as \(\text{info}(\varphi)\), and to express an issue which is resolved precisely by those propositions that are in \(\llbracket \varphi \rrbracket\). It is assumed that if a proposition \(p\) resolves a certain issue, then any stronger proposition \(q \subset p\) resolves that issue as well. Thus, \(\llbracket \varphi \rrbracket\) is always downward closed: if it contains a proposition \(p\) it also contains any \(q \subset p\). Furthermore, it is assumed that the inconsistent proposition, \(\emptyset\), resolves any issue. Thus, \(\llbracket \varphi \rrbracket\) always contains \(\emptyset\) and is therefore always non-empty. Taken together, sentence meanings in inquisitive semantics are thus defined as non-empty, downward closed sets of propositions.

In some cases the issue expressed by a sentence \(\varphi\) is trivial, in the sense that it is already resolved by the information conveyed by \(\varphi\) itself. This occurs if \(\text{info}(\varphi) \in \llbracket \varphi \rrbracket\). A sentence
\( \varphi \) is called inquisitive just in case the issue it expresses is non-trivial, i.e., just in case \( \text{info}(\varphi) \not\in \llbracket \varphi \rrbracket \).

Finally, the semantic alternatives associated with a sentence \( \varphi \) are those propositions that contain precisely enough information to resolve the issue expressed by \( \varphi \). Technically, these are the maximal elements of \( \llbracket \varphi \rrbracket \):

\[
\text{alt}(\varphi) := \{ p \in \llbracket \varphi \rrbracket \mid \text{there is no } q \in \llbracket \varphi \rrbracket \text{ such that } p \subset q \}
\]

Note that this characterization of alternatives entails that one alternative can never be properly contained in another; otherwise it could not be a maximal element of \( \llbracket \varphi \rrbracket \). Also note that if \( \varphi \) is non-inquisitive, it is always associated with a unique alternative, namely \( \text{info}(\varphi) \). Vice versa, if \( \varphi \) generates multiple alternatives, then it cannot be the case that \( \text{info}(\varphi) \in \llbracket \varphi \rrbracket \), which means that \( \varphi \) must be inquisitive.\(^9\)

### 5.2 Semantic assumptions

Following Coppock and Brochhagen (2013b), we assume that at least sentences generate multiple alternatives. However, following a suggestion made by Schwarz (2016a) in his critique of C&B, we analyse at least \( n \) as generating just two alternatives, \([n]\) and \([n+1,...)\). For instance:

\[
\text{alt}(\text{Bill ate at least four apples}) = \{ [4], [5,...) \}
\]

Note that this brings the analysis more in line with Büiring’s intuition that at least \( n \) is semantically equivalent to a disjunction \( n \text{ or more} \).\(^10\)

As for more than \( n \), we follow C&B in assuming that it generates just one semantic alternative:

\[
\text{alt}(\text{Bill ate more than three apples}) = \{ [4,...) \}
\]

The meaning for at least \( n \) and more than \( n \) given here should, of course, be obtained from a general analysis of the two modifiers, one that allows us to analyze them in combination with arguments other than numerals as well. Building on Solt (2011) and Coppock (2016),

---

\(^9\)We should note here that there are several perspectives one can take on the connection between inquisitiveness, a semantic notion, and the communicative effects of sentences when uttered in discourse. The perspective assumed here, in the spirit of Groenendijk (2009) and Coppock and Brochhagen (2013b), is that even if a sentence is inquisitive, i.e., even if it semantically expresses a non-trivial issue, a speaker who utters this sentence in discourse does not necessarily raise this issue. In particular, she does not necessarily request a response that addresses the issue. Under this perspective, it is possible to assume that a disjunctive declarative sentence like \textit{John ate two or three apples} (with falling intonation) is inquisitive, just like the corresponding interrogative \textit{Did John eat two apples, or three?}. One could say that the former is used to make an assertion and the latter to ask a question, and that in making an assertion, speakers do not raise the issue that the uttered sentence expresses (perhaps their assertion still ‘evokes’ the issue, but the effect is weaker than in the case of a question). Another perspective that one could take (see, e.g., Ciardelli et al., 2015; Farkas and Roelofsen, 2017) is that the issue expressed by a sentence is always raised when the sentence is uttered in discourse. Under this perspective, it does not make sense to treat a disjunctive declarative as being inquisitive, on a par with the corresponding disjunctive question. This perspective allows for a more economical interface between semantics and discourse pragmatics, but is not directly compatible with the present proposal.

\(^10\)Since we assume that the alternatives for at least \( n \) are never nested, the analysis we are proposing here can be formulated in the basic inquisitive semantics framework. By contrast, the account of Coppock and Brochhagen (2013b) had to be formulated in a version of inquisitive semantics that allows for nested alternatives. As discussed in Ciardelli et al. (2016), this has a number of problematic repercussions.
we assume that an expression of the form *at least* $\varphi$ is interpreted relative to a context $c$ providing (i) a *comparison class*, which is a set of sentence meanings $\Phi$ (in the inquisitive semantics sense) including the meaning of the prejacent $[\varphi]$, and (ii) a *pragmatic strength ordering* $\triangleright_c$, which is a partial order on $\Phi$, possibly but not necessarily coinciding with entailment. Relative to such a context $c$, *at least* $\varphi$ is interpreted as follows:

\[ [\text{at least } \varphi] = \text{exh}_c[\varphi] \cup \text{stronger}_c[\varphi] \]

where:

- $\text{exh}_c[\varphi] = \{ p \in [\varphi] \mid \text{there is no non-empty } q \in [\varphi] \text{ such that } \Phi_q \subset \Phi_p \}$
- $\Phi_p = \{ P \in \Phi \mid p \text{ is compatible with } P \} \quad \text{(and similarly for } \Phi_q)$
- $p$ is compatible with $P \iff p \cap \text{info}(P) \neq \emptyset$
- $\text{stronger}_c[\varphi] = \bigcup \{ P \in \Phi \mid P \triangleright_c [\varphi] \}$

In words, $[\text{at least } \varphi]$ is obtained by taking the union of two sets of propositions. The first set is the exhaustive strengthening of $[\varphi]$ relative to $c$, $\text{exh}_c[\varphi]$. We assume that exhaustification amounts to minimization of compatibility with other elements in the given comparison class. Thus, $\text{exh}_c[\varphi]$ is the set of propositions $p$ in $[\varphi]$ that are compatible with a minimal set of elements in $\Phi$. In other words, it is impossible to find a consistent proposition $q \in [\varphi]$ that is compatible with a strict subset of the elements of $\Phi$ that $p$ is compatible with.

For instance, if $\Phi = \{ [[\text{Bill ate } n \text{ apples}] \mid n \in \mathbb{N}] \}$, assuming a one-sided semantics of bare numerals, and if the pragmatic strength ordering corresponds to entailment, then $\text{exh}_c[[\text{Bill ate four apples}]]$ is the set of propositions $p$ such that (i) $p \in [[\text{Bill ate four apples}]]$, i.e., every world in $p$ is one in which Bill ate four or more apples, and (ii) $p$ is not compatible with $\text{info}(\text{Bill ate } n \text{ apples})$ for any $n > 4$. Thus, as expected, $\text{exh}_c[[\text{Bill ate four apples}]]$ amounts to $[[\text{Bill ate only four apples}]]$.

The second set of propositions, which we denote as $\text{stronger}_c[\varphi]$, is simply obtained by taking the union of all the elements of the contextual comparison class $\Phi$ that are pragmatically stronger than $[\varphi]$. Thus, in the above example, $\text{stronger}_c[[\text{Bill ate four apples}]] = \bigcup \{ [[\text{Bill ate } n \text{ apples}] \mid n > 4] \}$ which, under the one-sided semantic interpretation of numerals assumed here, amounts to $[[\text{Bill ate five apples}]]$. Thus, as desired, we derive that $[[\text{Bill ate at least four apples}]] = [[\text{Bill ate only four apples}]] \cup [[\text{Bill ate five apples}]]$.\(^{11}\)

\(^{11}\)For ease of exposition, we have assumed above that the prejacent of *at least* is a sentential expression. However, in most cases the prejacent of *at least* is most naturally taken to be a quantifier. For instance, the structure of "at least Alice and Bob came" is naturally taken to be [[at least Alice and Bob][came]], and not [at least][Alice and Bob came]. Fortunately, it is not hard to adapt the treatment described above to fit this more realistic assumption. In inquisitive semantics, the semantic value of a sentence is a set of propositions, i.e., an object of type $(st, t)$, abbreviated as $T$. A property is an object of type $eT$, i.e., a function mapping individuals to sets of propositions. Finally, a quantifier is an object of type $(eT, T)$, i.e., a function mapping properties to sets of propositions. We can then analyze *at least* as a modifier of quantifiers, i.e., as denoting a function from quantifiers to quantifiers. Assuming that for any quantifier $Q$, the context of use supplies a comparison class $\Phi_c$ consisting of multiple quantifiers including $Q$ itself, as well as a pragmatic strength ordering $\triangleright_c$ over $\Phi_c$, the entry for *at least* can be formulated as follows:

\[ [\text{at least}] = \lambda Q. (eT, T). \lambda A. T. \text{exh}_c(Q, A) \cup \text{stronger}_c(Q, A) \]

where:
We assume that sentences of the form more than ϕ are also interpreted relative to a contextually given, partially ordered comparison class. Only in this case, the meaning of the sentence simply amounts to $\text{stronger}_c[ϕ]$.

(30) $[\text{more than } ϕ] = \text{stronger}_c[ϕ]$  

To briefly illustrate how the general treatment of at least fares outside the domain of modified numerals, consider the following example:

(31) At least Ann left.

Suppose that the contextually given comparison class is as in (32), and that the strength ordering amounts to entailment:

(32) $Φ = \{[\text{Ann left}], [\text{Bill left}], [\text{Chris left}], [\text{Ann and Bill left}], [\text{Bill and Chris left}], [\text{Chris and Ann left}], [\text{Ann and Bill and Chris left}]\}$

Then we get that:

(33) a. $\text{exh}_c[(31)] = [\text{Only Ann left}]$  
b. $\text{stronger}_c[(31)] = \bigcup\{[\text{Ann and Chris left}], [\text{Ann and Bill left}]\}$

This means that:

(34) $\text{alt}((31)) = \{w \mid \text{Only Ann left in } w\}, \{w \mid \text{Ann and Bill left in } w\}, \{w \mid \text{Ann and Chris left in } w\}$

5.3 Pragmatic assumptions

5.3.1 Inquisitive sincerity

Following Coppock and Brochhagen (2013b) and earlier work on inquisitive pragmatics (Groenendijk and Roelofsen, 2009), we assume a maxim of inquisitive sincerity: if a speaker utters an inquisitive sentence ϕ, the issue that the sentence expresses should not already be resolved in the speaker’s information state.

(ii) a. $\text{stronger}_c((Q, A)) = \bigcup\{Q'(A) \mid Q' \succ_c Q\}$  
b. $\text{exh}_c((Q, A)) = \{p \in Q(A) \mid \text{there is no non-empty } q \in Q(A) \text{ such that } Φ_q^A \subset Φ_p\}$  
c. $Φ_p = \{Q'(A) \mid Q' \in Φ_c \text{ and } p \text{ is compatible with } Q'(A)\}$

Something that we have not made explicit here, since it is orthogonal to our concerns and would lead us quite far astray, is that the comparison class $Φ$ should be constrained by the focus structure of ϕ. Coppock and Brochhagen (2013b) capture this by letting the stronger alternatives come from the QUD. We are not taking the alternatives directly from the QUD here, so the link will have to be made indirectly.

Note that $\{w \mid \text{Ann and Bill and Chris left in } w\}$ is also an element of $[(31)]$, but it is not a maximal element, so it is not one of the semantic alternatives associated with the sentence.

The original formulation of the inquisitive sincerity maxim in Groenendijk and Roelofsen (2009) makes reference to the common ground: “If a speaker utters a sentence ϕ that is inquisitive w.r.t. the common ground, then ϕ should be inquisitive w.r.t. the speaker’s information state as well.” For our current purposes this qualification is not necessary. Note also that Coppock and Brochhagen operate with a
quisitive sentence \( \varphi \) implicates that the speaker’s information state is located in the set \( \text{inq-sincerity}(\varphi) \), defined as follows:

\[
\text{inq-sincerity}(\varphi) := \{ s \mid \text{if } \varphi \text{ is inquisitive, then } s \not\in [\varphi] \}
\]

### 5.3.2 Quality (informative sincerity)

As usual, we assume the Gricean maxim of Quality: if a speaker utters a sentence \( \varphi \), her information state \( s \) should support the informative content of \( \varphi \): \( s \subseteq \text{info}(\varphi) \). By this maxim, an utterance of \( \varphi \) implicates that the speaker’s information state is located in the set \( \text{quality}(\varphi) \) which is defined as follows:

\[
\text{quality}(\varphi) := \{ s \mid s \subseteq \text{info}(\varphi) \}
\]

Note that Quality requires speakers to be sincere, on a par with the maxim of inquisitive sincerity. Only, while inquisitive sincerity pertains to the issue expressed by the sentence that is uttered, Quality pertains to its informative content. Groenendijk and Roelofsen (2009) emphasise this similarity by referring to Quality as the maxim of informative sincerity. We will also sometimes use this terminology below. However, it is important for our purposes to stress that, while the maxims of inquisitive and informative sincerity are similar in nature, they are two independent maxims. In particular, in certain contexts speakers may be expected to comply with one of them but not necessarily with the other. The importance of this independence will become clear below.

### 5.3.3 Quantity

Following Schwarz (2016b) and many others, we assume that the maxim of quantity is concerned with alternative expressions that the speaker could have used. However, only expressions that are relevant to the question under discussion should be taken into consideration. Thus, unlike Schwarz, we distinguish lexical formal alternatives from contextual formal alternatives. The set of lexical formal alternatives for a sentence \( \varphi \) is denoted as \( A_{\varphi} \).

Following Schwarz, we assume that the lexical formal alternatives for at least \( n \) are \( \{ \text{at least } m \mid m \in \mathbb{N} \} \) and \( \{ \text{only } m \mid m \in \mathbb{N} \} \), and similarly, the lexical formal alternatives for more than \( n \) are \( \{ \text{more than } m \mid m \in \mathbb{N} \} \) and \( \{ \text{only } m \mid m \in \mathbb{N} \} \):

\[
\text{Lexically determined formal alternatives}
\]

\[\begin{align*}
\text{a. } \text{at least } n & : \{ \text{at least } m \mid m \in \mathbb{N} \} \cup \{ \text{only } m \mid m \in \mathbb{N} \} \\
\text{b. } \text{more than } n & : \{ \text{more than } m \mid m \in \mathbb{N} \} \cup \{ \text{only } m \mid m \in \mathbb{N} \}
\end{align*}\]

The set of contextual formal alternatives for a sentence \( \varphi \) relative to a question under discussion \( Q \), denoted \( A_{\varphi,Q} \), contains all and only those lexical formal alternatives for \( \varphi \) that are wholly relevant to \( Q \):

\[
A_{\varphi,Q} := \{ \psi \in A_{\varphi} \mid \psi \text{ is wholly relevant to } Q \}
\]

What does it mean for \( \psi \) to be wholly relevant to \( Q \)? Recall that the semantic alternatives in \( \text{alt}(Q) \) are propositions that contain precisely enough information to resolve the issue stronger sincerity maxim, which they call the maxim of interactive sincerity. On their account this is needed because the predictions that inquisitive sincerity delivers are too weak. On the present account, inquisitive sincerity delivers the right predictions, and interactive sincerity would do so as well; indeed, in the basic inquisitive semantics framework, interactive sincerity and inquisitive sincerity are equivalent.
expressed by \( Q \). They can be thought of, then, as wholly relevant, complete resolutions of \( Q \), i.e., complete resolutions that do not include any potentially irrelevant information. Similarly, any union of two or more such alternatives can be thought of as a wholly relevant, partial resolution of \( Q \). Thus, we say that \( \psi \) is wholly relevant to \( Q \) if and only if \( \text{info}(\psi) \) coincides with the union of a set of semantic alternatives in \( \text{alt}(Q) \).

\[(39) \quad \psi \text{ is wholly relevant to } Q \text{ iff } \text{info}(\psi) = \bigcup Q' \text{ for some } Q' \subseteq \text{alt}(Q).\]

We now turn to specifying exactly how quantity implicatures are derived. To facilitate direct comparison with Schwarz’s account, we will stay as close as possible to his innocent exclusion-based recipe. The only difference is that we restrict attention to contextual formal alternatives rather than always taking all lexical formal alternatives into account.

First, we compute primary quantity implicatures: if a speaker utters a sentence \( \varphi \), then her information state must not support the informative content of any contextual pragmatic alternative that would have been more informative than \( \varphi \). Let \( A_{\varphi,Q}^c \) denote the set of contextual formal alternatives that would have been more informative than \( \varphi \):

\[(40) \quad A_{\varphi,Q}^c = \{ \psi \in A_{\varphi,Q} | \text{info}(\psi) \subset \text{info}(\varphi) \} \]

Then, an utterance of \( \varphi \) in the context of a question \( Q \) implicates that the speaker’s information state is included in the set \( \text{quantity1}(\varphi, Q) \) which is defined as follows:

\[(41) \quad \text{quantity1}(\varphi, Q) = \{ s | \text{for all } \psi \in A_{\varphi,Q}^c : s \not\in \text{quality}(\psi) \} \]

Next, we compute secondary quantity implicatures. The recipe for doing so is the same as on Schwarz’s proposal, except that we now take \( Q \) into consideration. That is, we identify all formal alternatives \( \psi \) in \( A_{\varphi,Q}^c \) that are innocently excludable w.r.t. \( \varphi \) and \( Q \). This holds just in case for every subset \( A' \) of \( A_{\varphi,Q}^c \), if it is possible to find an information state in \( \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \) that rejects every sentence in \( A' \), then it is also possible to find a state in \( \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \) that rejects every sentence in \( A' \) as well as \( \psi \).

Secondary quantity implicatures are based on the assumption that the speaker’s state rejects all innocently excludable alternatives. Thus, we assume that an utterance of \( \varphi \) implicates that the speaker’s state is located in the set \( \text{quantity2}(\varphi, Q) \), defined as follows:

\[(42) \quad \text{quantity2}(\varphi, Q) = \{ s | \text{s rejects any } \psi \in A_{\varphi,Q}^c \text{ innocently excludable w.r.t. } \varphi \text{ and } Q \} \]

If inquisitive sincerity, informative sincerity / quality, and quantity implicatures are all taken into account, what a hearer will conclude is that the speaker’s information state must be in the set \( \text{cooperative}(\varphi, Q) \) defined as follows:

\[(43) \quad \text{cooperative}(\varphi, Q) = \text{inq-sincerity}(\varphi) \cap \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \cap \text{quantity2}(\varphi, Q) \]

5.3.4 Variance in robustness and strength of implicatures

We assume that, depending on the specific context at hand, speakers may be expected to comply more strictly with certain maxims than with others. This means that hearers,}

---

15In determining which formal alternatives are innocently excludable, one may also restrict oneself to the class of information states \( \text{sincerity}(\varphi) \cap \text{quantity1}(\varphi, Q) \) rather than the broader class \( \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \). This would not affect the predictions of our account, at least not in the empirical domain we are considering here.

16Though this assumption is sometimes left implicit, we take it to be commonplace. For instance, Grice (1975, p.27) writes: “It is obvious that the observance of some of these maxims is a matter of less urgency

---
when assessing a particular utterance in a given context, may sometimes give little weight to certain classes of implicatures—in some settings they may, for instance, take informative sincerity and quantity implicatures into account, but not inquisitive sincerity implicatures.

We therefore expect that implicatures which arise through two independent maxims will be more robust than implicatures that arise through only one of these two maxims. For, even if one of the two maxims is disregarded in a certain context, these implicatures will still arise through the other maxim.

We also expect that the effects of implicatures that arise through two independent maxims, when measured across a population of language users and across different conversational settings, will be stronger than those of implicatures that arise through only one of the maxims. To see this, let M1 and M2 be two independent maxims. Then, for any language user u and any conversational setting s, the chance that x takes either M1 or M2 (or both) to be active in s will be higher than the chance that she will take M1 to be active in s, and also higher than the chance that she will take M2 to be active in s. Thus, the cumulated effects of implicatures that arise through both M1 and M2 will be stronger than those of implicatures that arise only through M1, or only through M2.

Given these considerations, our pragmatic theory does not only make predictions as to whether or not a certain kind of implicature arises in a certain context, but also predictions concerning the comparative robustness and strength of different kinds of implicatures. These predictions are summarised in the following linking hypothesis.

(44) Linking hypothesis

Implicatures which arise through two independent maxims are more robust than implicatures that arise through only one of these maxims, and their effects, when measured across different language users and conversational settings, will be stronger.

As we will see in the next section, this hypothesis is crucial in explaining our experimental findings. In particular, it will account for the robustness and strength of the ignorance implicatures of at least in responses to a how many question, since on our account these implicatures arise through two independent maxims—inquisitive sincerity and quantity.

6 Predictions

We now show that our account derives the following results. When at least is used in the context of a how many question, an ignorance implicature is derived through both inquisitive sincerity and quantity. In the context of a polar question, it is derived through inquisitive sincerity only. For more than in how many contexts, ignorance is derived through quantity only. For more than in polar question contexts, ignorance is not derived through either route. These predictions are in line with the distinctions in robustness and strength of ignorance implicatures found in our experiments.

6.1 Predictions in the context of a how many question

Suppose the question under discussion is the how many question in (45a), which we associate with the set of semantic alternatives in (45b).

than is the observance of others; a man who has expressed himself with undue prolixity would, in general, be open to milder comment than would a man who has said something he believes to be false”.

33
Figure 10: At least 3 in a how many context. Contextual alternatives are depicted below the horizontal line.

(45)  
  a.  
  b.  

Superlative modifiers: First, let us consider the sentence in (46a), involving the superlative modifier at least. The semantic alternatives associated with this sentence on our account are given in (46b).

(46)  
  a.  
  b.  

Informative sincerity (quality) requires that $s \subseteq \text{info}(\varphi)$, that is, $s \subseteq [3, \ldots)$. On the other hand, since $\varphi$ is inquisitive, inquisitive sincerity requires that $s \not\subseteq [\varphi]$; that is, it requires $s$ not to be included in either of the semantic alternatives for $\varphi$; in other words, the speaker should not believe that the number of apples was exactly three, nor should she believe that the number is larger than three. So, from inquisitive sincerity we already derive that the speaker should be ignorant about the number of apples that John ate.

(47)  

Next, consider quantity implicatures. We have assumed that the lexical formal alternatives for $\varphi$ are sentences of the form $\psi_n = \text{John ate at least } n \text{ apples}$ or of the form $\chi_n = \text{John ate only } n \text{ apples}$, for $n \in \mathbb{N}$. All of these sentences are relevant for the question $Q$, and therefore they qualify as contextual pragmatic alternatives. Thus, $A_{\varphi,Q}^\subseteq$ consists of the sentences $\psi_n$ with $n > 3$, as well as $\chi_n$ with $n \geq 3$. This is depicted in Figure 10.

Primary quantity implicatures require that the speaker could not have uttered any of these sentences without violating informative sincerity. This means that $s \not\subseteq [3]$ and $s \not\subseteq [4, \ldots)$. That is, the speaker should consider it possible that John ate exactly three apples, and she should also consider it possible that John ate more than three apples. Thus, precisely the same ignorance implicature that was derived through inquisitive sincerity is also derived as a primary quantity implicature.
(48) \( \text{quantity1}(\varphi) = \text{inq-sincerity}(\varphi) \)

Finally, we will show that no contextual formal alternative in \( A_{\varphi, Q}^{\leq} \) is innocently excludable, which means that no secondary quantity implicatures arise. Consider for instance the formal alternative \( \psi_5 \), \textit{John ate at least 5 apples}. To see that this alternative is not innocently excludable, consider the set \( A' = \{\chi_4\} \), where \( \chi_4 = \text{John ate only 4 apples} \). Rejecting \( \chi_4 \) is consistent with the quality and primary quantity implicatures, since some information states in \( \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \) reject it (as a witness, take the state \([3] \cup [5])\). But rejecting \( \chi_4 \) as well as our candidate alternative \( \psi_5 \) (\textit{John ate at least 5 apples}) is not consistent with the quality and primary quantity implicatures; no information state in \( \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \) rejects both \( \chi_4 \) and \( \psi_5 \). In other words, rejecting \( \psi_5 \) forces acceptance of \( \chi_4 \), which means that the former is not innocently excludable. Similar reasoning holds for all the other at least alternatives \( \psi_n \) with \( n > 3 \). To see this, consider the set \( A'_n = \{\chi_4, \ldots, \chi_{n-1}\} \) (in particular, take \( A'_4 = \emptyset \)). Some information states in \( \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \) reject all elements of \( A'_n \): if \( n = 4 \), this holds trivially, as \( A'_4 = \emptyset \); if \( n > 4 \), we can take as a witness the state \([3] \cup [n] \). However, no information state in \( \text{quality}(\varphi) \cap \text{quantity1}(\varphi, Q) \) rejects all elements of \( A'_n \) in addition to \( \psi_n \).

Similarly, to show that each 'only alternative' \( \chi_n \) is not innocently excludable we can take \( A'_n = \emptyset \) if \( n = 3 \), and \( A'_n = \{\chi_m | m \geq 4, m \neq n\} \) if \( n > 3 \).

Since none of the alternatives is innocently excludable, no secondary quantity implicatures arise, or more precisely, nothing new is concluded about the state of the speaker by drawing secondary quantity implicatures.

In sum, we have that:

(49) \( \text{cooperative}(\varphi, Q) = \{s | s \subseteq [3, \ldots) \text{ and } s \not\subseteq [3] \text{ and } s \not\subseteq [4, \ldots)\} \)

Thus, we predict that, from an utterance of (46a) in the context of the \textit{how many} question in (45a), an ignorance implicature is drawn, and no upper bounding implicature. Importantly, the relevant ignorance implicature is not just that the speaker does not know exactly how many apples John ate, but also that the speaker does not know whether John ate exactly three apples or more. And since this implicature is derived both through inquisitive sincerity and through quantity, it is predicted to be more robust than ignorance implicatures that arise through only one of these maxims.\(^\text{17}\)

**Comparative modifiers:** Now consider (50a). We have assumed that this sentence is associated with a unique semantic alternative, given in (50b).

(50) a. \( \varphi : \text{John ate more than two apples} \).

b. \( \text{alt}(\varphi) = [3, \ldots) \)

Let us compute what implicatures are predicted for \( \varphi \) in the context of \( Q \). First, informative sincerity requires that \( s \subseteq \text{info}(\varphi) \). That is, the speaker should believe that John ate at least three apples.

Since \( \varphi \) is not inquisitive, the maxim of inquisitive sincerity does not place any additional constraints on the speaker’s state. In particular, no ignorance implicature arises

\(^\text{17}\) Although modified numerals of the form ‘\textit{n or more}’ are not within the immediate scope of the present paper, it is worth mentioning that, given an inquisitive treatment of disjunction and assuming that the first disjunct (i.e., \( n \)) is exhausted to obviate redundancy (Chierchia et al., 2012; Katzir and Singh, 2013; Ciardelli and Roelofsen, 2017), our account makes exactly the same predictions for ‘\textit{n or more}’ as for ‘at least \textit{n}’, which seems a desirable result.
through inquisitive sincerity in this case.

Now let us turn to quantity implicatures. The set of lexical formal alternatives for \( \varphi \) consists of all sentences of the form \( \psi_n = \text{John ate more than } n \text{ apples} \) and all sentences \( \chi_n = \text{John ate only } n \text{ apples} \), for any natural number \( n \). The situation is depicted in Figure 11. All of these formal alternatives are relevant to the question \( Q \) we are considering, and therefore qualify as contextual pragmatic alternatives. Thus, we get the following primary quantity implicatures:

\[
(51) \quad \text{quantity1}(\varphi, Q) = \{ s \mid s \not\subseteq [3] \text{ and } s \not\subseteq [4, \ldots] \}
\]

This means that we derive ignorance as a primary quantity implicature: the speaker should not know whether John ate exactly three apples or more.

Finally, consider secondary quantity implicatures. Note that none of the formal alternatives \( \psi \in A^c_{\varphi, Q} \) is innocently excludable, for reasons parallel to those given for \( \varphi_{\text{at least}} \) above. This means that secondary quantity implicatures do not arise, which leads to the following overall result:

\[
(52) \quad \text{cooperative}(\varphi, Q) = \{ s \mid s \subseteq [3, \ldots] \text{ and } s \not\subseteq [3] \text{ and } s \not\subseteq [4, \ldots] \}
\]

Thus, for the \textit{more than} sentence (50a) in the context of the \textit{how many} question in (45a) we predict an ignorance implicature and no upper bounding implicature. While the relevant ignorance inference is exactly the same that was derived above for (46a), there is a crucial difference between the two cases: in the case of (50a) the inference is only derived as a quantity implicature, while in the case of (46a) it was derived both through quantity and through inquisitive sincerity. Given the Linking Hypothesis in (44), this provides an account for the experimental results pertaining to \textit{how many} contexts reported in Section 4: we found that in the case of \textit{at least}, ignorance implicatures arise very robustly, across different experimental settings, while in the case of \textit{more than} they arise in some experimental settings, but not in all. We propose that those settings in which \textit{more than} does not give rise to ignorance implicatures (i.e., those in which participants had to judge the appropriateness of a given utterance while having access to the speaker’s information
state) are ones in which quantity implicatures are not given much weight, while sincerity implicatures, both informative and inquisitive ones, are taken into full consideration.

6.2 Predictions in the context of a polar question

We now turn to polar question contexts. What is special about such contexts in our view is that the formal alternatives that are lexically associated with a given expression, are often not relevant w.r.t. the question under discussion, and therefore do not play a role in the computation of quantity implicatures.

Suppose that John’s diet prescribes that he eat at most two apples per day, and does not prescribe anything else. Consider the polar question in (53a). Given our contextual assumptions, the semantic alternatives associated with this question are the ones given in (53b).

(53) a. Did John stick to his diet today?
   b. \( Q = \{[0, 2], [3, \ldots] \} \)

Superlative modifiers: First, let us consider the response to \( Q \) in (54a) below, involving the superlative modifier \( \text{at least} \). The semantic alternatives associated with this sentence on our account are given in (54b).

(54) a. \( \varphi : \text{No, John ate at least three apples.} \)
   b. \( \text{alt}(\varphi) = \{[3], [4, \ldots] \} \)

We have seen above that for this sentence, ignorance is derived as an inquisitive sincerity implicature; since such implicatures are QUD-independent on our account, this inference still goes through in the present setting.

As for quantity implicatures, note that none of the lexically determined formal alternatives for \( \varphi \) is relevant w.r.t. \( Q \). Thus, the situation is as depicted in Figure 12. This means that quantity implicatures do not arise, and we get the following overall result:

(55) \( \text{cooperative}(\varphi, Q) = \{s \mid s \subseteq [3, \ldots] \text{ and } s \not\subseteq [3] \text{ and } s \not\subseteq [4, \ldots] \} \)

Thus, an ignorance inference is derived in this case as well, but only through inquisitive sincerity—not through quantity, unlike in the case of \( \text{at least} \) sentences in the context of a \textit{how many} question. Given our Linking Hypothesis in (44), this provides an explanation of the finding that, the effects of ignorance inferences triggered by \( \text{at least} \) are, when considered across different experimental settings, stronger in the context of a \textit{how many} question than in the context of a polar question. Recall that this contrast was found in Experiments 1 and 2, and a similar tendency was observed (though not found significant) in Experiment 3.

Also recall that the results of Experiment 2 were in fact compatible with the assumption that, in this specific experimental setting, \( \text{at least} \) sentences did not give rise to ignorance inferences at all in response to polar questions. If this is indeed the case, then it could be accounted for by assuming that in performing the task that participants were given in this experiment (i.e., judging whether or not the speaker had full knowledge about the number of relevant items based on what she said), they did not give much weight to the maxim of inquisitive sincerity.
Comparative modifiers: Now consider (56a). This sentence is associated with a unique semantic alternative, given in (56b).

(56) a. $\varphi$: No, John ate more than two apples.
    b. $\text{alt}(\varphi) = \{[3, \ldots]\}$

Again, none of the lexically determined formal alternatives for $\varphi$ are relevant w.r.t. the given polar QUD. Thus, the situation is as depicted in Figure 13. This means that no quantity implicatures arise. In particular, no ignorance inference is derived, unlike in the context of a how many question. This accounts for the fact that in none of our experiments, ignorance implicatures were detected for more than sentences in response to polar questions.18

6.3 Further prediction: Mendia’s generalization

Before concluding, we would like to highlight one further prediction of the account, concerning an empirical generalization recently put forward by Mendia (2016).

To track the reasoning that led to this generalization, first consider sentence (57) below. We have seen that such sentences typically implicate that the speaker is uncertain as to whether exactly two students completed the quiz or more.

(57) At least two students completed the quiz.

However, Mendia (2016) argues that this inference depends crucially on the fact that, for numerals, the relevant comparison class is linearly ordered. In cases which involve a partially ordered comparison class, such as (58), the speaker might in fact know that Ann and Bill are not the only students who completed the quiz.

(58) At least Ann and Bill completed the quiz.

18In addition, notice that, due to the absence of contextually relevant formal alternatives, we also correctly predict the lack of upper bounding implicatures for both (54a) and (56a).
To further support this claim, Mendia provides data showing that the answer in (59b) is regarded considerably more acceptable than the one in (60b).

(59)  
   a. Who completed the quiz?
   b. I don’t remember, at least Ann and Bill, but not only them.

(60)  
   a. How many students completed the quiz?
   b. #I don’t remember, at least two, but not only two.

This leads Mendia to propose the following generalization:

(61)  
   a. When the relevant comparison class forms a partial order, an *at least* sentence need not convey speaker ignorance as to whether the exhaustified prejacent holds;
   b. When the relevant comparison class is a total order, an *at least* sentence does always convey speaker ignorance with respect to the exhaustified prejacent.

This generalization is predicted by our account. For the *at least Ann and Bill* part of (59b) we get the semantic value in Figure 14 (where *a* and *b* stand for Ann and Bill, respectively, and *c* and *d* for two additional people in the domain of discourse). Under the pragmatic assumptions we have laid out, the sentence can be felicitously uttered by a speaker as long as her information state is consistent with multiple semantic alternatives. This requirement can be satisfied even if the speaker’s information state excludes the exhaustified prejacent alternative, {*a*, *b*}, because there are multiple additional alternatives, due to the fact that the comparison class does not form a total linear order in this case.

On the other hand, the first conjunct in (60b) only generates two semantic alternatives, namely ‘exactly two’ and ‘more than two’, because the comparison class does form a total linear order here. Therefore, ignorance with respect to both of these alternatives is predicted.

7 Conclusion

We set out to address two related questions in this paper: (i) At the empirical level, when exactly do superlative and comparative modified numerals give rise to ignorance implicatures? And (ii) at a theoretical level, what are the sources of such implicatures?

In the literature, there is disagreement with respect to both questions. Empirically, the received view is that superlative modifiers like *at least* trigger ignorance implicatures, while comparative modifiers like *more than* don’t (e.g., Geurts and Nouwen, 2007). Recently,
however, a number of different views have been put forward (Coppock and Brochhagen, 2013b; Mayr and Meyer, 2014; Westera and Brasoveanu, 2014). Theoretically, the main divide is between approaches that derive ignorance inferences of modified numerals from the maxim of quantity (e.g., Schwarz, 2016b; Mayr and Meyer, 2014; Westera and Brasoveanu, 2014) and an alternative approach which derives such inferences from the maxim of inquisitive sincerity (Coppock and Brochhagen, 2013b).

The experiments reported in the present paper show that the empirical picture is more complex than has been assumed previously:

1. The ignorance implicatures of *at least* in *how many* contexts are strong and robust: they were observed across all experimental settings.

2. The ignorance implicatures of (a) *at least* in polar contexts and (b) *more than* in *how many* contexts are less robust: they were not detected in all experimental settings. Moreover, the experimental settings in which the former were detected were not the same as those in which the latter were detected. Finally, ignorance inferences are usually weaker in these cases (and certainly never stronger) than those of *at least* in *how many* contexts.

3. We did not detect any ignorance implicatures triggered by *more than* in polar contexts.

We proposed to account for these findings by merging the main ideas from the quantity-based approaches and the inquisitive sincerity-based approach. In the resulting theory, ignorance inferences may arise both through quantity and through inquisitive sincerity, depending on the construction and the context at hand. Moreover, depending on the task that participants are asked to perform in a specific experimental setting, the maxim of quantity and the maxim of inquisitive sincerity may be given more or less importance. In particular, when performing a task in which the speaker’s perspective is taken, i.e., when judging the appropriateness of a given utterance given a specification of the speaker’s knowledge state, we propose that the maxim of quantity is given relatively little importance, while when performing a task in which the hearer’s perspective is taken, i.e., when inferring what the speaker’s knowledge state is based on what she said in a given context, the maxim of inquisitive sincerity does not play a major role. We have argued that this dual-route approach to ignorance implicatures is in a better position to account for our experimental findings than any of the single-route approaches developed previously.

References


Alrenga, P. (2016). *At least* and *at most*: Scalar focus operators in context. Manuscript, Boston University.


Mayr, C. and M.-C. Meyer (2014). More than at least. Slides presented at the *Two days at least* workshop, Utrecht.


