Natural Language Semantics
https://doi.org/10.1007/s11050-024-09226-3

The ups and downs of ignorance

Marco Degano[1] · Paul Marty[2] · Sonia Ramotowska[1] · Maria Aloni[1] · Richard Breheny[3] · Jacopo Romoli[4] · Yasutada Sudo[3]

Accepted: 10 December 2024

Abstract

Plain disjunctive sentences, such as *The mystery box contains a blue ball or a yellow ball*, typically imply that the speaker does not know which of the two disjuncts is true. This is known as an IGNORANCE inference. We can distinguish between two aspects of this inference: the negated universal upper bound part (i.e., the speaker is uncertain about each disjunct), which we call UNCERTAINTY, and the existential lower bound part (i.e., the speaker considers each disjunct possible), which we call POSSIBILITY. In the traditional approach, UNCERTAINTY is derived as a primary implicature, from which POSSIBILITY follows. In this paper, we report on two experiments using a sentence-picture verification task based on the mystery box paradigm that challenge the traditional implicature approach. Our findings show that POSSIBILITY can arise without UNCERTAINTY, and we thus call for a reevaluation of the traditional view of disjunction and IGNORANCE inferences. Our experimental findings are related to similar results involving disjunction in embedded contexts and pave the way for alternative theories that can account for the observed patterns of inference derivation in a unified fashion. We discuss how recent implicature and non-implicature theories can account for the derivation of existential lower bound inferences without the presence of negated universal upper bound inferences.

Keywords Experimental semantics & pragmatics · Disjunction · Implicature ignorance inferences

1 Introduction

It is well-known that sentences with a plain disjunction like (1) typically give rise to IGNORANCE and EXCLUSIVITY inferences, indicated in (1-a) and (1-b):

(1) Sue went to the park or the cinema. $A \vee B$

a. ⇝ *The speaker doesn't know which of the two is true* IGNORANCE

b. ⇝ *Sue didn't go to the park and to the cinema* EXCLUSIVITY

Extended author information available on the last page of the article

Published online: 20 January 2025

In this work, we mainly focus on IGNORANCE inferences, though we also consider the role of EXCLUSIVITY when relevant. A natural way to think about IGNORANCE is as follows: a speaker s is ignorant about A when s does not believe that A is the case and does not believe that $\neg A$ is the case. In other words, both $\neg A$ and A are live possibilities in the speaker's doxastic state (Hintikka 1962; Gazdar 1979). We can thus distinguish between two aspects of an ignorance inference of a plain disjunction like (1): the upper bound UNCERTAINTY part in (2-a) and the lower bound POSSIBILITY part in (2-b), where $\Box_s$ and $\Diamond_s$ quantify universally and existentially over the speaker's doxastic state.

(2) $\Diamond_s A \wedge \neg\Box_s A \wedge \Diamond_s B \wedge \neg\Box_s B$ IGNORANCE INFERENCE
The speaker doesn't know which of the two is true.
  a. $\neg\Box_s A \wedge \neg\Box_s B$ UNCERTAINTY
  The speaker is not certain that Sue went to the park and the speaker is not certain that Sue went to the cinema.
  b. $\Diamond_s A \wedge \Diamond_s B$ POSSIBILITY
  The speaker deems possible that Sue went to the park and the speaker deems possible that Sue went to the cinema.

The traditional approach to ignorance inferences derives UNCERTAINTY as an implicature of (1) through Gricean quantity reasoning about each individual disjunct. From this, POSSIBILITY follows, given also the quality assumption that the speaker believes (1) (Sauerland 2004b; Fox 2007, among others). As such, this approach predicts that POSSIBILITY *cannot* arise in the absence of UNCERTAINTY.[1]

Recent non-implicature proposals, however, suggest that POSSIBILITY should be derived independently of UNCERTAINTY, either through a presupposition (Goldstein 2019) or as so-called neglect-zero pragmatic inference (Aloni 2022). According to these theories, a sentence like (1) can give rise to POSSIBILITY in the absence of UNCERTAINTY. Consequently, these theories, unlike the traditional implicature approach, predict that a disjunctive sentence like (1) should be judged as felicitous when the speaker is certain about one of the disjuncts, while only considering the other possible.

In this study, we report on three experiments testing these divergent predictions. All experiments consisted of a sentence-picture verification task based on the mystery box paradigm (Noveck 2001; Marty et al. 2023, 2024). The results of our experiments show that POSSIBILITY can indeed arise without UNCERTAINTY. Specifically, we found that participants judged sentences where UNCERTAINTY inferences were false as good as true controls, while more than half of them judged the sentence false when POSSIBILITY inferences were also false. These results pose a challenge for the traditional approach, which cannot predict POSSIBILITY in the absence of UNCERTAINTY.

Our experiments focus on cases of plain disjunctive sentences. Sentences involving disjunction in embedded contexts give rise to similar inference patterns and have

[1] Grammatical versions of the implicature approach make the same prediction, but they derive the corresponding inferences by means of a covert exhaustivity operator, interacting with a covert doxastic operator (Meyer 2013; Buccola and Haida 2019, among others).

sparked a similar debate (Crnič et al. 2015; Marty et al. 2023). We examine the relationship between our findings and previous experimental research on embedded disjunction. In particular, we interpret our experimental results from three different perspectives.

Firstly, we explore how a recent grammatical account (Bar-Lev and Fox 2023), originally developed to capture the behavior of disjunction in embedded contexts, can be extended to deal with plain disjunctive sentences and what the consequences of this extension are. Secondly, our experimental results are in line with the recent non-implicature accounts of POSSIBILITY mentioned above (Goldstein 2019; Aloni 2022). We focus on the account of Aloni (2022) and discuss how it derives POSSIBILITY. Thirdly, we comment on the possible linking hypotheses behind our experiments and consider how our results fare with game-theoretic approaches to disjunction and modal inferences (Franke 2009). We conclude with potential follow-ups and a discussion on the role of different experimental tasks and protocols for the same types of inferences, emphasizing the importance of taking a holistic experimental approach when studying pragmatic phenomena.

The rest of the paper is structured as follows. Section 2 summarizes the traditional approach and its predictions. Section 3 contains an overview of the experiments, followed by a report of Experiment 1, Experiment 2 and Experiment 3 in Sects. 4, 5 and 6, respectively. Section 7 discusses the results of our experiments, focusing on the connection with distributive inferences and the grammatical approach in Sect. 7.1, implicature accounts of disjunction in Sect. 7.2, non-implicature accounts of disjunction in Sect. 7.3, and game-theoretic approaches in Sect. 7.4. Section 8 concludes.

2 The traditional approach and its predictions

The traditional Gricean approach (Grice 1975, 1989) derives IGNORANCE and EXCLUSIVITY inferences based on a general process of conversation that takes into account the information available to speakers and the assumption that language users behave cooperatively. The most relevant principles underlying these derivations are the maxim of quantity, which roughly says that speakers should convey all and only the most informative statement given what they know; and the maxim of quality, which roughly says that speakers should be always truthful.

These maxims can be made formal in different ways and a number of accounts have been proposed stemming from the original Gricean insights (Horn 1972; Gazdar 1979; Levinson 1983; Gamut 1991, among others). In what follows, we rely on the following formalization of inference derivation:

(3) For all relevant alternatives ψ of an utterance ϕ:

$$\text{if } \psi \models \phi \text{ and } \phi \not\models \psi, \text{ then } \neg\Box_s(\psi)$$

In words, when speakers choose to make a weaker statement (ϕ) instead of a stronger alternative (ψ), it implies that they do not believe that the stronger statement is true ($\neg\Box_s\psi$). The principle above refers to the notion of *relevant* alternatives. An important question for this approach is to determine what counts as an alternative, and what ultimately counts as a relevant alternative.

We assume, as proposed in Sauerland (2004b) and maintained in subsequent literature, that a disjunction has also the individual disjuncts as alternatives. And these alternatives will end up being relevant in most cases.

(4)
$$\left\{\begin{array}{ll}\text{Sue went to the park or the cinema.} & A \vee B\\ \text{Sue went to the park.} & A\\ \text{Sue went to the cinema.} & B\\ \text{Sue went to the park and to the cinema.} & A \wedge B\end{array}\right\}$$

The Gricean account just outlined provides a systematic way to derive both IGNORANCE and EXCLUSIVITY inferences associated with a plain disjunction. As previously stated, IGNORANCE inferences are comprised of two elements: UNCERTAINTY and POSSIBILITY. The standard procedure is to derive UNCERTAINTY first, and then use quality to derive POSSIBILITY.[2]

For instance, $\neg\Box_s A$ would be derived as in (5). A similar process with the alternative B and with the alternative $A \wedge B$ would give us $\neg\Box_s B$ and $\neg\Box_s(A \wedge B)$, respectively. Together, these are known as the primary implicatures of a plain disjunctive sentence.

(5) Derivation of UNCERTAINTY ($\neg\Box_s A$)

a. Assertion: $A \vee B$
b. A is a relevant alternative to $A \vee B$
c. A is stronger than $A \vee B$
d. From (5-b) and (5-c) above together with (3), $\neg\Box_s A$

POSSIBILITY is then derived from UNCERTAINTY together with the Quality assumption that speakers believe $A \vee B$. As a result, this method of deriving POSSIBILITY makes this inference dependent on UNCERTAINTY.

(6) Derivation of POSSIBILITY

a. Assertion: $A \vee B$
b. Quality: $\Box_s(A \vee B)$
c. UNCERTAINTY: $\neg\Box_s A \wedge \neg\Box_s B$
d. From (6-b) and (6-c) above, $\Diamond_s A \wedge \Diamond_s B$

As for the derivation of the EXCLUSIVITY inference in (1-b), a similar argument to the derivation in (5) for the alternative $A \wedge B$ yields $\neg\Box_s(A \wedge B)$. However, to obtain the EXCLUSIVITY inference $\Box_s\neg(A \wedge B)$, the standard approach relies on an opinionatedness principle of the form $\Box_s\phi \vee \Box_s\neg\phi$. This principle, also known as *competence assumption* or *epistemic step*, states that the speaker knows the relevant facts (Sauerland 2004b, 2005; Geurts 2010; Chierchia et al. 2012). Applying this principle to $\neg\Box_s(A \wedge B)$ gives us the (secondary) implicature $\Box_s\neg(A \wedge B)$.

[2] Another option is to close the alternatives under negation, so that the negative counterparts of the positive alternatives count as relevant (Fox 2007). Even with this approach, the negative alternatives, responsible for POSSIBILITY, are present due to the positive alternatives, responsible for UNCERTAINTY.

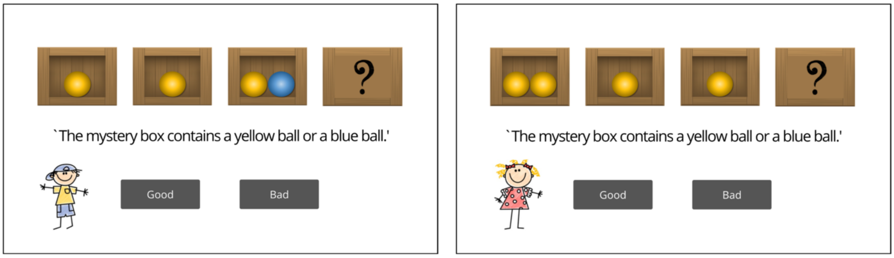


Fig. 1 Example items illustrating the items' layout and task in Experiments 1, 2 and 3. These examples are instances of the TARGET-1 (on the left) and TARGET-2 (on the right) conditions in Experiment 2. UNCERTAINTY is false in both the TARGET-1 and TARGET-2 conditions while POSSIBILITY is false only in the TARGET-2 conditions

It is important to note here is that the derivation of POSSIBILITY in (6) does not rely on the opinionatedness principle just outlined. In fact, Sauerland (2004b) claims that individual disjuncts cannot be subject to opinionatedness, as generating inferences of the form $\Box_s \neg \psi$ from $\neg \Box_s \psi$ would, in combination with the Quality assumption about the assertion, contradict previously obtained implicatures (i.e. $\neg \Box_s \psi'$).[3]

3 Overview of the experiments

As discussed in the previous section, the derivation of POSSIBILITY relies on UNCERTAINTY according to the standard approach. In the following, we report on three experiments testing this hypothesis, which we outline here.

All experiments were built upon Marty et al.'s (2024) elaboration of Noveck's (2001) mystery box paradigm (see also Marty et al. 2023). In all experiments, participants were presented with sentence-picture items like those in Fig. 1. Each item depicted a set of four boxes, three open and one covered, displayed just above an utterance produced by one of two characters. Participants were instructed that the characters could see what's inside the first three boxes but not what's inside the covered one, the so-called *mystery box*. They were also instructed that the characters had been taught the rule that the mystery box always had the same contents as one of the three open boxes (see Sect. 4.3 and the supplementary material available online on

[3]Relatedly, we note that one of the driving factors behind the richer set of alternatives in (4) was the observation that UNCERTAINTY and POSSIBILITY, unlike standard scalar implicatures, are not cancelable (Sauerland 2004a,b). However, we observe that while POSSIBILITY is indeed not cancelable, UNCERTAINTY seems to pattern with EXCLUSIVITY in terms of its cancelability:

(i) a. #I saw the Queen or the Princess of the Netherlands at the airport.
In fact, it cannot be that I saw the Queen.
b. I saw the Queen or the Princess of the Netherlands at the airport.
In fact, I saw both.
c. I saw the Queen or the Princess of the Netherlands at the airport.
In fact, I am pretty sure I saw the Queen and possibly the Princess.

the journal website for the experimental instructions). The task was to decide if the character's utterance was right given the information available to them and the rule that they had learned about how the mystery box works. Participants reported their response by clicking on one of two response buttons, labeled 'Good' and 'Bad'. Example (7) illustrates the disjunctive test sentences studied in the experiments and its related inferences.

(7) 'The mystery box contains a yellow ball or a blue ball.' $\Box_s(A \vee B)$

a. $\neg\Box_s A \wedge \neg\Box_s B$ UNCERTAINTY

b. $\Diamond_s A \wedge \Diamond_s B$ POSSIBILITY

The two critical conditions were TARGET-1 and TARGET-2. In the TARGET-1 conditions, test sentences were paired with pictures that make UNCERTAINTY false, but POSSIBILITY true; in the TARGET-2 conditions, they were paired with pictures that make both these inferences false. We hypothesized that, if POSSIBILITY is derived from UNCERTAINTY, no difference in participant's responses should be observed between both target conditions. On the other hand, if POSSIBILITY is derived independently from UNCERTAINTY, then participants should reject the test sentences to a greater extent in the TARGET-2 than in the TARGET-1 conditions.

4 Experiment 1

4.1 Participants

101 native speakers of English participated in this study (mean age 39 yrs; 51 female). Participants were recruited online through Prolific (https://www.prolific.co; see Palan and Schitter for an overview) using a suitable set of prescreening criteria (first language: English; nationality: UK/US; country of birth: UK/US; minimum approval rate: $\geq 90\%$). Participants were paid £1.20, and median completion time was about 8 minutes (hourly rate: £9/hr). All participants gave written informed consent. Data were collected and stored in accordance with the provisions of Data Protection Act 2018. The study was approved by the Research Ethics Committee at University College London and at University of Amsterdam.

4.2 Materials and design

The experiment was based on the materials and method from Marty et al. (2024, Exps. 4-6) (see also Marty et al. 2023). Each item involved a sentence displayed just below a set of boxes and right above the picture of one of two characters, as exemplified in Fig. 1. Sentences were constructed using the sentence frames in Table 1. The color adjectives used in the sentences, indicated by the [A] and [B] terms in Table 1, were picked at random from a list of four color terms—*yellow*, *blue*, *green* and *gray*—with replacement across items. Each sentence was placed inside simple quotation marks to show direct speech and to make it explicit that the sentence was uttered by whichever character was depicted on the item.

Table 1 Schematic description of the sentences tested in Experiment 1, where [A] and [B] are placeholders for different color adjectives; for a more concrete illustration, you may read [A] as *yellow* and [B] as *blue*

TEST		'The mystery box contains a [A] ball or a [B] ball.'
CONTROL	C1	'The mystery box contains a [A] ball.'
	C2	'The mystery box does not contain a [A] ball.'
	C3	'The mystery box contains a [A] ball and a [B] ball.'
	C4	'The mystery box does not contain a [A] ball or a [B] ball.'

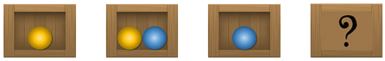
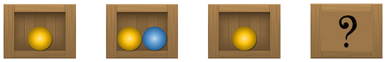
Every item displayed a set of four boxes horizontally arranged, each of which was made of three open boxes, containing one or two balls, and a covered one, marked with the symbol '?' and referred to as *the mystery box*. The position of the boxes in the row was pseudo-randomly assigned so that the mystery box always appeared in the rightmost position. The contents of the open boxes in each quadruplet were experimentally manipulated to create different picture types corresponding to the experimental conditions of the study. The test sentences under investigation were paired with five different picture types, which are described in Table 2: the colors of A-balls and B-balls depicted in the open boxes always matched the [A] and [B] color terms used in the sentences (e.g., yellow and blue), whereas the colors of the C-balls and D-balls were randomly chosen from our list of color terms by excluding the color(s) of the matching balls (e.g., green and gray).

Target pictures were designed to make UNCERTAINTY always false, but POSSIBILITY either true or false. On the TARGET-1 pictures, each of the three open boxes contained an A-ball and at least one of them also contained a B-ball, making one of the UNCERTAINTY inferences of the test sentences false, but their POSSIBILITY inferences true. TARGET-2 pictures were obtained from the TARGET-1 pictures by replacing the B-ball(s) with balls of a non-matching color (i.e., a C-ball or a D-ball), thus making one of the UNCERTAINTY and one of the POSSIBILITY inferences false. Different variants of the TARGET-1 and TARGET-2 pictures were constructed by varying the number of matching B-balls for the former (1 B-ball vs. 2 B-balls) and by varying both the number and color of non-matching balls for the latter (1 C-ball vs. 2 C-balls vs. 1 C-ball and 1 D-ball). For the purposes of experimental design, variants of the TARGET-1 and TARGET-2 pictures were treated as sub-conditions of the TARGET-1 and TARGET-2 conditions.[4]

FALSE, TRUE-EXCL and TRUE-ADHOC pictures were control pictures, each of which served a different experimental purpose. FALSE pictures were designed to provide a clear baseline for rejection. On these pictures, one of the open boxes contained an A-ball, while the other two contained balls of a non-matching color, making the test sentences unambiguously false. TRUE-EXCL and TRUE-ADHOC pictures were designed to probe for the strength of two other inference types that the test sentences could give rise to: (i) "not-both" exclusivity implicatures, e.g., *the mystery box does not contain both a [A] ball and a [B] ball* and (ii) ad hoc exhaustivity inferences, e.g., *the mystery box does not contain a [C] ball or a [D] ball*. Crucially, the derivation

[4] As reported in Sect. 4.5, no contrast in participants' responses was found between the variants of the TARGET-1 pictures, nor between those of the TARGET-2 pictures, allowing us to aggregate the responses to the TARGET-1 and the TARGET-2 trials across sub-conditions without loss of information.

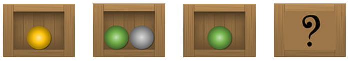
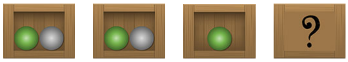
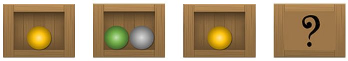
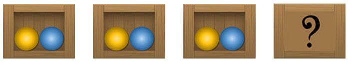
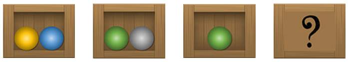
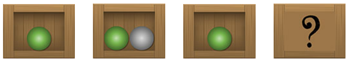
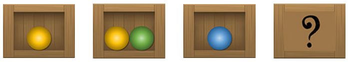
Table 2 Schematic description and illustration of the picture types paired with the test sentences in Experiment 1. The color of the A-balls and B-balls always matched the color adjectives used in the sentence (e.g., *yellow* and *blue*) while the color of the C-balls and D-balls never did (e.g., *green* and *gray*). Picture types are illustrated here using the following color assignment: A=yellow, B=blue, C=green and D=gray

Condition		Example picture				
TRUE-ADHOC			A	AC	B	?
TRUE-EXCL			A	AB	B	?
TARGET-1	i.		A	AB	A	?
	ii.		A	AB	AB	?
TARGET-2	i.		A	AC	A	?
	ii.		A	AC	AC	?
	iii.		A	AC	AD	?
FALSE			A	CD	C	?

Test sentence: *'The mystery box contains a yellow ball or a blue ball.'* $(A \vee B)$

of such inferences can affect participants' responses in the target conditions: the first one makes the test sentences false in the TARGET-1 conditions, where one or more boxes contain two matching balls, while the second makes them false in the TARGET-2 conditions, where one or more boxes contain a non-matching ball. The goal of the TRUE-EXCL and TRUE-ADHOC conditions was to assess the extent to which participants derived these inferences so as to factor out their potential effects from the comparison of primary interest between TARGET-1 and TARGET-2 conditions. On the TRUE-EXCL pictures, one of the open boxes contained both an A-ball and a B-ball, as on the TARGET-1 pictures. These pictures made the test sentences true unless an exclusivity inference is derived (i.e., *the mystery box contains a yellow ball or a blue, but not both*). On the TRUE-ADHOC pictures, one of the open boxes contained an A-ball and a C-ball, as on the TARGET-2 pictures. These pictures made the test sentence

Table 3 Schematic description and illustration of the picture types paired with the control sentences in Experiment 1. Picture types are illustrated using the same color assignment as before

Sentence	Condition	Example picture
C1: A	GOOD	 A A A ?
	BAD	 A CD C ?
C2: $\neg A$	GOOD	 CD CD C ?
	BAD	 A CD A ?
C3: $A \wedge B$	GOOD	 AB AB AB ?
	BAD	 AB CD C ?
C4: $\neg(A \vee B)$	GOOD	 C CD C ?
	BAD	 A CD B ?

true unless an ad hoc exhaustivity inference is derived (i.e., *the mystery box contains a yellow ball or a blue, and nothing else*).

In addition to the test sentences, there were four different types of control sentences: two positive sentences (C1 and C3) and two negative ones (C2 and C4), involving either one color adjective (C1 and C2) or two (C3 and C4). Each of these sentences were paired with pictures that made them either clearly true (GOOD conditions) or clearly false (BAD conditions), as described and illustrated in Table 3. These control items were added to the experiment to identify low-effort responses as well as to control for certain low-level response strategies in the data treatment. Specifically, we worried that some participants may perform the task superficially, simply by checking whether or not the colors mentioned in the sentence match those of the balls depicted on the pictures. We reasoned that, if a participant follows such a strategy, then they should perform relatively poorly on these items, especially in those involving negative sentences (C2 and C4).

Pairing the test and control sentences with the relevant picture types gave rise to 5 test and 8 control conditions. Each control condition was instantiated 3 times, giving rise to 24 control trials. As for the test conditions, the TARGET-1, TARGET-2 and FALSE conditions were instantiated 6 times each, and the TRUE-EXCL and TRUE-ADHOC conditions 3 times each, giving rise to 24 test trials. Instances of the TARGET-1 and TARGET-2 conditions were evenly distributed across their respective sub-conditions. Thus, each survey included 48 trials in total.

4.3 Procedure

The experiment was run as an online survey using Gorilla Experiment Builder (Anwyl-Irvine et al. 2020). In the instructions, participants were introduced to two characters, Sam and Mia, and they were presented with a short cover story. The cover story went in substance as follows (see the supplementary material available online on the journal website for details): Sam and Mia are looking at quadruplets of boxes containing balls of various colors. Each time, they can only see what's inside the first three boxes. However, they have been taught the rule that the fourth box, known as "the mystery box", always has the same contents as one of the three open boxes and, therefore, they can make certain inferences about what's inside the mystery box. Participants were then shown two examples illustrating what the characters can and cannot infer thanks to this rule. Participants were told that the characters would be presented with many quadruplets, each of which would be followed by an utterance from either Sam or Mia about what the mystery box contains, and that their task was to decide if this utterance was right given the information available to the characters and the rule that they learned. They were instructed to click on 'Good' if they consider that the characters got it right and, otherwise, to click on 'Bad'.

Following the instructions, participants completed a short training devised to consolidate their understanding of the cover story. The training phase included one instance of each control condition, hence 8 trials. During this phase, participants received feedback on the accuracy of their responses, together with a short explanation as to why the character got it right or wrong (e.g., *Sam got it right: since every open box contains a yellow ball, one can be certain that the mystery box does too*). After the training, the study continued with a block of 48 experimental trials. Trials were presented in random order, with a 1000 ms inter-stimulus interval. Participants reported their responses by clicking with the mouse one of two response buttons labeled 'Good' and 'Bad', respectively. The position of the response buttons (i.e., on the left or on the right) was counterbalanced across participants. Items remained on the screen until participants gave their response.

4.4 Data availability

The stimuli, data and analysis code associated with Experiment 1 are available open access on the OSF platform at https://osf.io/4ut2c/.

4.5 Data treatment and analysis

Data treatment and analysis were carried out in the R statistical environment (R Core Team 2021) using the Hmisc (Harrell 2021), lme4 (Bates et al. 2015), car (Fox and Weisberg 2019) and sjPlot (Lüdecke 2023) packages for the R statistics program.

Responses from 15 participants were removed prior to analysis because their performance in the control trials did not reach the pre-established threshold of 80% accuracy.[5] The performance of the remaining subjects (n=86) was uniformly high both in the BAD conditions (M=94.7%, 95% CI[93.2, 95.9]) and the GOOD conditions (M=95.9%, 95% CI[94.5, 96.9]), with little variation among sentence types (all Ms>92.2% across all control conditions). Responses to the TARGET-1 and TARGET-2 trials were next inspected to check for potential discrepancies among the different variants of the TARGET-1 and TARGET-2 pictures (see Table 2). For each target condition, we fitted a mixed effect logistic regression model, predicting responses from the fixed effect of picture sub-type. In both cases, the maximal converging model only included a random intercept for participants. Both models were compared to a null model missing the fixed effect of interest. Neither model was significantly different from the null model (TARGET-1: $\chi^2(1) = 1.52$, $p = .21$; TARGET-2: $\chi^2(2) = 4.05$, $p = .13$), meaning that picture sub-type had no detectable effect on participants' judgments in the target conditions. Based on these results, responses to the TARGET-1 and TARGET-2 trials were aggregated across sub-conditions for the purpose of the main analysis.

We fitted generalized linear mixed-effects regression models (binomial distribution, logit link function), predicting responses to the target sentences in the test trials from the fixed effect of CONDITION (5 levels, dummy coded).[6] To comprehensively examine all contrasts, we run two separate models: one with target-1 and the other with target-2 as the reference level for the fixed effect of CONDITION. Both models included a random slope for CONDITION within *Subject* and a correlated random intercept. Here, and in the following models, we included the maximal random effects structure justified by the experimental design (Barr et al. 2013). We then simplified the model as necessary to ensure proper convergence.

[5] The responses of the excluded participants were investigated to identify the possible sources of their low performance. We found little evidence for the low-level response strategy mentioned in Sect. 4.2. Rather, we found a discrepancy between the GOOD and BAD conditions (86% vs. 55% accuracy), showing that these participants generally (incorrectly) accepted the control sentences in the BAD conditions. Given the contents of the pictures in these conditions, it suggests that these participants generally considered an utterance as right as long as it was true of at least one box, that is, if it was describing a possible outcome (rather than a certainty). Whether this bias toward acceptance stems from an application of the Charity Principle or simply from a poor understanding of the instructions, it justifies further the conservative criterion we used for inclusion in the analysis.

[6] Another approach would have been to fit two models—one with and one without an independent effect of possibility—and compare their fit to the data. We explored this type of analysis in the scripts available in the OSF repository. Another potential method is to conduct pairwise comparisons between conditions. Although this would reduce the number of contrasts, it complicates the estimation of random effects, as participants could receive different random intercepts for each pair of conditions. Fitting the model to the entire dataset addresses this issue. We thank a reviewer for these insights.

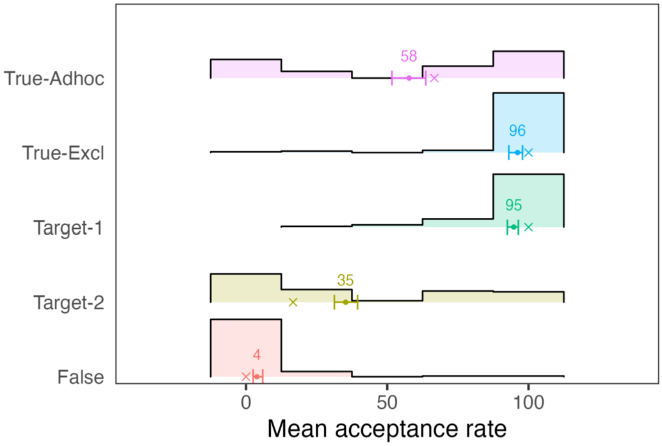


Fig. 2 Mean acceptance rate (i.e., proportion of 'Good' responses) to test sentences by experimental condition in Experiment 1. For each condition, the distribution of by-participant mean rates is visualized by a histogram, the grand mean by a dot with its rounded value on top and the 95% CI around it, and the median by a cross

Table 4 Summary of statistical models outputs and analyses. The R pseudo-code for the `glmer` models is `Response` ∼ 1 + `Condition` + (1 + `Condition|Subject`). The coding scheme is dummy coding on `Condition`. z-values are calculated based on the estimated coefficients and their standard errors. p-values are calculated using Wald tests. Confidence intervals are calculated using the Wald method. For random effects estimates, refer to the analysis script. The mid-line separates the model with TARGET-1 as the reference level from the one with TARGET-2 as the reference level. Conditional and Marginal R^2 values for both models are 0.96 and 0.59, respectively

	Estimate	95% CI	z-value	p-value
Intercept	5.07	[3.26, 6.88]	5.49	**<.001**
TARGET-1.TARGET-2	−6.81	[−8.75, −4.87]	−6.88	**<.001**
TARGET-1.TRUE-EXCL	5.36	[−0.61, 11.32]	1.76	0.078
TARGET-1.TRUE-ADHOC	−4.29	[−6.26, −2.31]	−4.25	**<.001**
TARGET-1.FALSE	−14.40	[−18.46, −10.36]	−6.97	**<.001**
Intercept	−1.74	[−2.80, −0.68]	−3.22	0.0013
TARGET-2.TARGET-1	6.81	[4.87, 8.75]	6.87	**<.001**
TARGET-2.TRUE-EXCL	12.20	[6.29, 18.04]	4.05	**<.001**
TARGET-2.TRUE-ADHOC	2.53	[1.54, 3.51]	5.01	**<.001**
TARGET-2.FALSE	−7.60	[−11.32, −3.87]	−4.00	**<.001**

4.6 Results

Figure 2 shows the mean acceptance rate (i.e., the proportion of 'Good' responses) for the test sentences as a function of the experimental condition. The outputs of the statistical models and analyses are summarized in Table 4.

Acceptance rates were uniformly high in the TARGET-1 (M=94.7%, 95% CI[92.4, 96.3]) and TRUE-EXCL conditions (M=96.1%, 95% CI[93.0, 97.8]), with no significant difference between the two ($z = 1.76$, $p = 0.078$). These results show that the participants in our study did not derive the UNCERTAINTY inferences or the exclusivity implicatures associated with the test sentences. By contrast, the mean acceptance rate for the TARGET-2 and TRUE-ADHOC conditions were somewhat intermediate (M=35.2%, 95% CI[31.2, 39.4] and M=57.7%, 95% CI[51.6, 63.6], respectively), between the high(est) rates observed in the TARGET-1 and TRUE-EXCL conditions and the low(est) rate observed in the FALSE conditions (M=3.8%, 95% CI[2.5, 5.9]), with the TARGET-2 conditions yielding significantly lower rates than the TRUE-ADHOC conditions ($z = 5.01$, $p < 0.001$).[7] The results from the TRUE-ADHOC conditions indicate that participants derived the ad hoc exhaustivity inferences associated with the test sentences to a noticeable extent. Since these inferences were false both in the TRUE-ADHOC and TARGET-2 conditions, it is possible that their derivation also affected to some extent participants' judgments in the TARGET-2 conditions. Thus, the absolute rates of acceptance for the TARGET-2 conditions in this experiment should be regarded with caution. Crucially, however, the contrast between the TARGET-2 and TRUE-ADHOC conditions show that, everything else being equal, participants rejected the test sentences significantly more often, when their POSSIBILITY inferences were also false. This contrast is explained if, in contrast to the UNCERTAINTY inferences, participants sometimes derived the POSSIBILITY inferences associated with the test sentences, possibly in addition to deriving ad hoc exhaustivity inferences.

4.7 Discussion

The results of this experiment show that participants uniformly accepted simple disjunctive sentences of the form 'A or B' in cases where UNCERTAINTY was false, but POSSIBILITY was true (TARGET-1 conditions), whereas they rejected these same sentences 65% of the time in cases where both UNCERTAINTY and POSSIBILITY were false (TARGET-2 conditions). Taken at face value, these findings go against the traditional approach to IGNORANCE inferences, which predicts POSSIBILITY inferences to arise from and, thus to be licensed by, the derivation of UNCERTAINTY inferences. They suggest instead that POSSIBILITY inferences remain available to speakers even in the absence of the associated UNCERTAINTY inferences and, consequently, that the former can be derived independently of the latter.

As we discussed, however, the present results also suggest that the derivation of POSSIBILITY inferences may not be the only driver behind the contrast between the TARGET-1 and TARGET-2 conditions and, specifically, behind the low acceptance rates observed in the TARGET-2 conditions. We also found that the test sentences

[7] A visual inspection of the histograms in Fig. 2 suggest that, unlike in the TARGET-1 conditions, subjects' responses in the TARGET-2 and TRUE-ADHOC conditions were not uniform. To verify this impression, we carried out a post hoc exploration of individual subjects' means by testing for unimodality of their distribution via Hartigans' dip tests in these three conditions, which indicated non-unimodality for TARGET-2 and TARGET-ADHOC but not for TARGET-1 (TARGET-1: $D = 0.04$, $p = 0.18$; TARGET-2: $D = 0.09$, $p < .001$; TRUE-ADHOC: $D = 0.14$, $p < .001$). For completeness, a Pearson correlation coefficient was computed to assess the linear relationship between subjects' mean rates in the TARGET-2 and TRUE-ADHOC conditions. The results show a moderate, positive correlation ($r(84) = 0.49$, 95% CI[0.32, 0.64], $p < .001$).

were rejected 42% of the time in the TRUE-ADHOC conditions, showing that participants also derived their ad hoc exhaustivity inferences (e.g., ‘A or B, and nothing else’) to a noticeable extent. Since deriving these inferences would also lead to rejecting the test sentences in the TARGET-2 conditions, the interpretation of the contrast between the TARGET-1 and TARGET-2 conditions is conditional on that of the further contrast between the TRUE-ADHOC and TARGET-2 conditions. For the time being, we propose that, everything else being equal, the fact that the TARGET-2 conditions gave rise to significantly lower acceptance rates than the TRUE-ADHOC conditions supports the idea that the results to the TARGET-2 conditions cannot be entirely explained by the putative effect of ad hoc exhaustivity inferences, that is, without also taking into account the falsity of the POSSIBILITY inferences in these conditions.

One potential worry with the interpretation above is that it relies on the assumption that the robustness and the frequency at which ad hoc exhaustivity inferences were derived remained relatively constant across experimental conditions. While this assumption is certainly not implausible, the present results offer no independent evidence that would support it. In particular, we cannot exclude the possibility that (i) ad hoc exhaustivity inferences were derived in the TARGET-2 conditions and (ii) these inferences were more robust, or simply derived more often in the TARGET-2 than in the TRUE-ADHOC conditions. As is easy to see, making the alternative assumptions in (i) and (ii) would allow one to similarly account for the contrasts at hand without any need to refer to POSSIBILITY inferences at all.

In the following section, we report on a follow-up experiment testing this alternative explanation of the data. For these purposes, the test items from Experiment 1 were minimally modified to nullify the effect of ad hoc exhaustivity inferences on participants’ judgments so as to obtain more pristine TARGET-2 conditions. The results of this experiment show that the key contrast between TARGET-1 and TARGET-2 reproduces once the potential effects of ad hoc exhaustivity inferences are factored out from the comparison, thus confirming our original interpretation of the data.

5 Experiment 2: factoring out ad hoc exhaustivity inferences

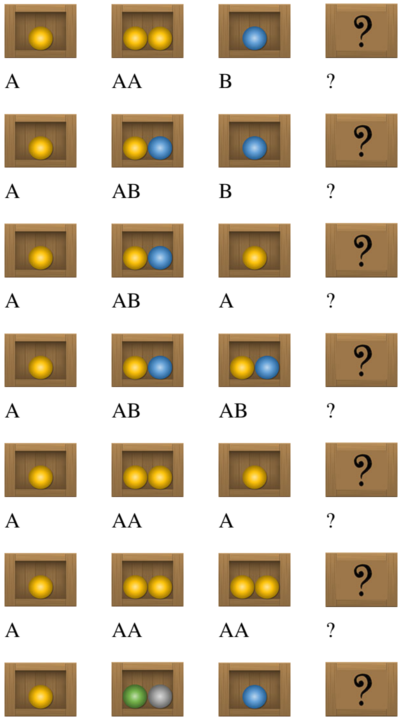
5.1 Participants

100 novel participants took part in this study (mean age 42 yrs; 50 female). Participants were recruited online through Prolific using the same prescreening criteria as in Experiment 1. Participants were paid £1.20, and median completion time was about 8 minutes (hourly rate: £9/hr). The consent and data collection procedures were the same as in Experiment 1.

5.2 Materials and design

The materials were the same as in Experiment 1, except for the TARGET-2, TRUE-ADHOC and FALSE pictures, which were minimally modified to address specific issues left open by the results of Experiment 1. For simplicity, the novel conditions resulting from these picture modifications were labelled TARGET-2*, TRUE* and

Table 5 Schematic description and illustration of the picture types paired with the test sentences in Experiment 2. The materials were the same as in Experiment 1, except for the pictures used in the novel TRUE*, TARGET-2* and FALSE* conditions. Picture types are illustrated using the same color assignment as before (A=yellow, B=blue, C=green and D=gray)



Condition		Example picture			
TRUE*		A	AA	B	?
TRUE-EXCL		A	AB	B	?
TARGET-1	i.	A	AB	A	?
	ii.	A	AB	AB	?
TARGET-2*	i.	A	AA	A	?
	ii.	A	AA	AA	?
FALSE*		A	CD	B	?

Test sentence: *'The mystery box contains a yellow ball or a blue ball.'* $(A \vee B)$

FALSE*, respectively. The picture types (and subtypes) used in the test trials of Experiment 2 are described and illustrated in Table 5.

First, the TARGET-2 pictures were modified to make the ad hoc exhaustivity inferences associated with the test sentences true, while still making both UNCERTAINTY and POSSIBILITY false. The motivation was to create a more pristine version of these conditions enabling us to factor out the potential effect of ad hoc exhaustivity inferences from the comparison of primary interest. The novel TARGET-2* pictures were obtained from the TARGET-2 pictures by replacing each ball of a non-matching color with an A-ball. As a result, all balls depicted on the TARGET-2* pictures were in effect A-balls. These modifications were applied to all variants of the TARGET-2 pictures resulting in two TARGET-2* sub-types, differing from each other only in terms of the number of A-balls they depicted. In the same vein, the TRUE-ADHOC pictures were modified to create genuinely "true" pictures. The novel TRUE* pictures were obtained from the TRUE-ADHOC pictures through the same procedure as above—i.e., by replacing any ball of a non-matching color with an A-ball—thus making the test sentences true irrespective of the pragmatic inferences they may give rise to. Therefore, participants were expected to robustly and uniformly accept the test sentences in these conditions. Finally, the FALSE pictures were modified to create more challenging false controls that can also be used to probe for low-effort responses in trials involving the test sentences (as opposed to other control sentences). On the novel FALSE* conditions, one of the open boxes contained an A-ball, another one contained a B-ball, while the last one contained balls of a non-matching color. Thus, if participants do not pay attention to all the open boxes in such cases (e.g., if they restrict their attention to boxes containing A-balls or B-balls), they should incorrectly accept these false controls and, therefore, we should be able to detect and quantify this behavior in our data.

The rest of the design was identical to that of Experiment 1 in all respects. Thus, in particular, Table 3 also stands as a summary of the GOOD and BAD control conditions used in Experiment 2. Each control condition was instantiated 3 times, giving rise to 24 control trials. The TARGET-1,TARGET-2* and FALSE* conditions were instantiated 6 times each, and the TRUE*and TRUE-EXCL conditions 3 times each, giving rise to 24 test trials. As before, instances of the 2 target conditions were evenly distributed across their respective sub-conditions. Thus, each survey included 48 trials in total, exactly as in Experiment 1.

5.3 Procedure

The procedure was identical to the one used in Experiment 1 (see Sect. 4.3 for details).

5.4 Data availability

The stimuli, data and analysis code associated with Experiment 2 are available open access on the OSF platform at https://osf.io/4ut2c/.

5.5 Data treatment and analysis

Data treatment was the same as in Experiment 1. Responses from 13 participants were removed because their performance in the control trials did not reach the pre-established threshold of 80% accuracy.[8] The performance of the remaining subjects (n=87) was uniformly high both in the BAD conditions (M=95.8%, 95% CI[94.4, 96.9]) and the GOOD conditions (M=97.0%, 95% CI[95.8, 97.9]), with little variation among sentence types (all Ms>93.6% across all control conditions). As before, for each target condition (i.e., TARGET-1 and TARGET-2*), we fitted a mixed effect logistic regression model, predicting responses from the fixed effect of picture sub-type (see Table 5). For both models, the maximal converging only included a random intercept for participants. Neither model was significantly different from the null

[8] For completeness, a descriptive analysis of the responses of the excluded participants was, here again, carried out (see fn. 5). As in Experiment 1, there was a discrepancy between the GOOD and BAD conditions across all control trials (81% vs. 57% accuracy). In this dataset, however, the relevant discrepancy was found to be mainly driven by higher rates of incorrect responses in the BAD conditions of the negative sentences C2 and C4 (48% and 33% accuracy, respectively). These errors are those predicted by the low-level response strategy that we mentioned in Sect. 4.2.

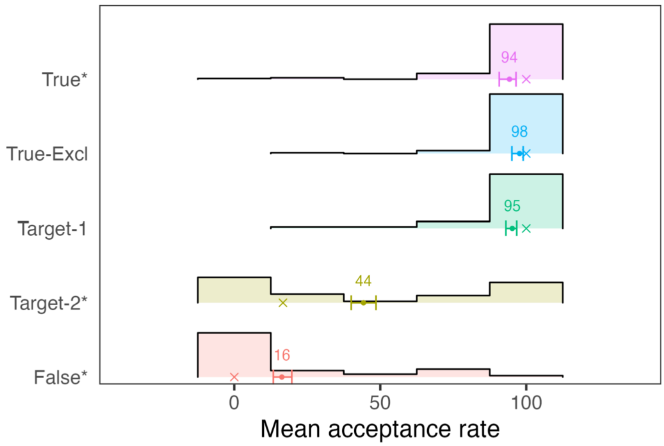


Fig. 3 Mean acceptance rate (i.e., proportion of 'Good' responses) to target sentences by experimental condition in Experiment 2. For each condition, the distribution of by-participant mean rates is visualized by a histogram, the grand mean by a dot with its rounded value on top and the 95% CI around it, and the median by a cross

model (TARGET-1: $\chi^2(1) = 0.54$, $p = .46$; TARGET-2*: $\chi^2(2) = 3.44$, $p = .06$). Responses to the TARGET-1 and TARGET-2* trials were thus aggregated across picture sub-conditions for the purpose of the main analysis.

We fitted generalized linear mixed-effects regression models (binomial distribution, logit link function), predicting responses to the target sentences in the test trials from the fixed effect of CONDITION. We employed successive difference coding on CONDITION (levels = [*True**, *True-Excl*, *Target-1*, *Target-2**, *False**]). This coding scheme creates contrasts by comparing each level of the categorical variable to the previous level, thus minimizing the number of comparisons. The model included a random slope of CONDITION within *Subject* and an uncorrelated random intercept. The random effects structure was simplified from Experiment 1 to avoid convergence or overfitting issues.

5.6 Results

Figure 3 shows the mean acceptance rate (i.e., the proportion of 'Good' responses) for the test sentences as a function of the experimental condition. The outputs of the statistical models and analyses are summarized in Table 6.

The mean acceptance rate in the TARGET-1 conditions was very high (M=95.2%, 95% CI[93.0, 96.7]), similar to those found in the TRUE-EXCL conditions (M=97.7%, 95% CI[95.0, 98.9], $z = -0.27$, $p = 0.79$) and in the novel TRUE* conditions (M=94.2%, 95% CI[90.7, 96.4], with no significant difference between TRUE* and TRUE-EXCL ($z = 0.24$, $p = 0.81$).[9] Thus, as in Experiment 1, there is no evidence

[9] Similar to what we observed in the TARGET-2 conditions of Experiment 1 (see fn. 7), the by-participant mean rates in the TARGET-2* conditions was not distributed unimodally ($D = 0.16$, $p < .001$), as shown in Fig. 3 by the presence of two peaks in their distribution. Unlike in Experiment 1, however, the distribution of the by-participant mean rates was otherwise unimodal in all other conditions.

Table 6 Summary of statistical models outputs and analyses. The R pseudo-code for the `glmer` model is `Response` ∼ 1 + `Condition` +(1|`Subject`) +(0 + `Condition`|`Subject`). The coding scheme is successive difference coding on `Condition`. z-values are calculated based on the estimated coefficients and their standard errors. p-values are calculated using Wald tests. Confidence intervals are calculated using the Wald method. For random effects estimates, refer to the analysis script. Conditional and Marginal R^2 values are 0.91 and 0.78, respectively

	Estimate	95% CI	z-value	p-value
Intercept	3.03	[1.41, 4.66]	3.65	**<.001**
TRUE*.TRUE-EXCL	0.59	[−4.26, 5.44]	0.24	0.81
TRUE-EXCL.TARGET-1	−0.69	[−5.77, 4.38]	−0.27	0.79
TARGET-1.TARGET-2*	−7.77	[−11.26, −4.28]	−4.37	**<.001**
TARGET-2*.FALSE*	−4.03	[−6.38, −1.68]	−3.36	**<.001**

that participants in this experiment ever derived the UNCERTAINTY inferences or exclusivity implicatures associated with the test sentences. By contrast, the mean acceptance rate in the novel TARGET-2* conditions was in the mid-range (M=44.2%, 95% CI[40.0, 48.5], between the high(est) rates observed in the TARGET-1, TRUE-EXCL and TRUE* conditions and the low(est) rate observed in the novel FALSE* conditions (M=16.2%, 95% CI[13.3, 19.6]). In particular, we found no significant difference between TARGET-2* and TARGET-1 ($z = -4.37$, $p<0.001$), and between TARGET-2 and FALSE* ($z = -3.36$, $p<0.001$). In sum, these results show that the key contrast observed in Experiment 1 remains once the potential effects of ad hoc exhaustivity inferences are factored out.

5.7 Discussion

We designed Experiment 2 as a minimal variant of Experiment 1 with the aim of factoring out the potential effects of ad hoc exhaustivity inferences from the comparison between the TARGET-1 and TARGET-2 conditions. The results that we obtained show that this aim was achieved and confirm the main findings from Experiment 1.

First, in line with our expectations, the novel TRUE* conditions were uniformly accepted with a mean acceptance rate of 94%, contra 57% for the TRUE-ADHOC conditions in Experiment 1. These results show that the picture modifications that we have made in Experiment 2 were successful at preventing conflicting ad hoc exhaustivity inferences from affecting participants' judgments. This, in turn, secures our interpretation of the comparison between the TARGET-1 and TARGET-2* conditions as assessing, specifically, the effect of false POSSIBILITY inferences on participants' judgments. Second, consistent with the results of Experiment 1, participants uniformly accepted the test sentences in the TARGET-1 conditions, whereas they robustly rejected them in the novel TARGET-2* conditions. Taken together, these results confirm the main findings of Experiment 1.

Finally, we note that the novel FALSE* conditions gave rise to slightly higher rates of errors compared to the original FALSE conditions (16% vs. 3%). These errors suggest that, on some occasions, some participants only verified whether or not balls of the matching colors were present somewhere on the pictures, i.e., disregarding whether or not these balls were present in every box. Assuming that this verification strategy was used at times, it may have magnified the contrast between the TARGET-1 and TARGET-2(*) conditions by pushing the participants further towards acceptance in the TARGET-1 conditions. The low frequency of these errors, however, indicates that the effect of this low-effort strategy was, at best, marginal among the participants and cannot account in any reasonable way for the large contrast in acceptance between the TARGET-1 and TARGET-2(*) conditions.

Before discussing the consequences of these results for current approaches to IGNORANCE inferences, let us consider yet another possible explanation for the contrasts observed between the TARGET-1 and TARGET-2(*) conditions, which was suggested to us by two anonymous reviewers.

This explanation starts from the observation that, in theory, sentences like *The mystery box contains a yellow ball or a blue ball* may allow for an interpretation on which the meaning of the first disjunct is strengthened locally. For expository purposes, we can think of this interpretation as the one delivered by embedding a covert instance of 'only' (notated $\mathcal{O}$ here) in the first disjunct, as illustrated in (8).[10]

(8) a. The mystery box contains ⟨only⟩ a yellow ball or a blue ball.
b. $\mathcal{O}(A) \vee B \equiv (A \wedge \neg B) \vee B$

On the traditional Gricean approach, the critical alternatives for calculating UNCERTAINTY and POSSIBILITY in this case are thus as follows:

(9) $\left\{\begin{array}{ll}\text{The mystery box contains only a yellow ball} & \mathcal{O}(A)\\ \text{The mystery box contains a blue ball} & B\end{array}\right\}$

Assuming that the stronger alternatives above are relevant, the Gricean account predicts for (8) the inferences in (10), by the same reasoning as the one presented in Sect. 2. As before, the UNCERTAINTY inferences in (10-a) are derived as primary implicatures by Quantity; the POSSIBILITY inferences in (10-b) are then derived from the former and from the Quality assumption that the speaker believes $\mathcal{O}(A) \vee B$.

(10) a. $\neg\Box_s\mathcal{O}(A) \wedge \neg\Box_s B$ UNCERTAINTY
b. $\Diamond_s\mathcal{O}(A) \wedge \Diamond_s B$ POSSIBILITY

The resulting inferences are thus somewhat different from those previously discussed for $A \vee B$. In particular, as far as UNCERTAINTY goes, in place of $\neg\Box_s A$ and $\neg\Box_s B$, we now have $\neg\Box_s\mathcal{O}(A)$ and $\neg\Box_s B$. Crucially, these UNCERTAINTY inferences, unlike the former, are compatible with the type of situations depicted in the TARGET-1 conditions. The reason is that, in these conditions, the color mentioned in the first disjunct (A) was consistently presented in all three visible boxes while the color mentioned in the second disjunct (B) was consistently presented in only one or two of them (see Tables 2 and 5 for examples). Thus, in these conditions, the speaker was always certain that A was true but uncertain as to whether B also was, consistent with the UNCERTAINTY (and POSSIBILITY) inferences in (10). This observation is impor-

[10] While this interpretation may be difficult to access upon introspection, it becomes more readily available if an additive particle is used in the second conjunct, e.g., 'The mystery box contains a yellow ball or *also/even* a blue ball'. For our purposes, we can think of such examples as variants of the example in (8).

Table 7 QUALITY, UNCERTAINTY and POSSIBILITY inferences for 'B or A', as predicted by the standard implicature approach, according to the two parses of interest. $\mathcal{O}$ indicates a covert instance of *only*. The last two columns show, for each parse, whether the predicted inferences are true (✓) or false (✗) in the TARGET conditions of Experiment 3

Parse	Inference		TARGET-1 A - AB - A(B)	TARGET-2 A - AA - A(A)
$B \vee A$	QUALITY:	$\Box_s(B \vee A)$	✓	✓
	UNCERTAINTY:	$\neg\Box_s B \wedge \neg\Box_s A$	✗	✗
	POSSIBILITY:	$\Diamond_s B \wedge \Diamond_s A$	✓	✗
$\mathcal{O}(B) \vee A$	QUALITY:	$\Box_s(\mathcal{O}(B) \vee A)$	✓	✓
	UNCERTAINTY:	$\neg\Box_s(\mathcal{O}(B) \wedge \neg\Box_s A$	✗	✗
	POSSIBILITY:	$\Diamond_s \mathcal{O}(B) \wedge \Diamond_s A$	✗	✗

tant because it means that, if participants in our studies parsed the test sentences as shown in (8), then the TARGET-1 conditions, as we designed them in Experiments 1 and 2, did not in fact discriminate between UNCERTAINTY and POSSIBILITY. Furthermore, note that on this alternative parse both UNCERTAINTY and POSSIBILITY would be true in the TARGET-1 conditions, whereas, as before, both inference types would be false in the TARGET-2 conditions. As a result, the availability of a parse along the lines of (8) would also account for the contrasts of interest, independently of whether or not POSSIBILITY lives on UNCERTAINTY.

In the following, we report on a third and final experiment testing this alternative explanation of the data. The TARGET conditions in this experiment were designed so that, if participants parse the test sentences as initially assumed, the key contrasts from Experiments 1 and 2 should reproduce once again, whereas if they parse these sentences as shown in (8), they should now reject them to the same extent in the TARGET-1 and TARGET-2 conditions and, therefore, the key contrasts should disappear. The results of this experiment align in every respect with those of Experiments 1 and 2, in support of the first hypothesis and against the second.

6 Experiment 3: controlling for alternative parses

Experiment 3 was designed to investigate the availability of the alternative parsing strategy that we described at the end of the previous section. For these purposes, we modified the test items in Experiment 2 by reversing the order of the disjuncts in the test sentences, changing it from 'A or B' to 'B or A'. These sentences were paired with the same picture types as before, thus resulting in novel TARGET conditions in which the color mentioned in the *second* disjunct (as opposed to the first one) was always represented in all three visible boxes. As illustrated in Table 7, these modifications allowed us to discriminate between the two parses of interest, namely $B \vee A$ and $\mathcal{O}(B) \vee A$, on the basis of the POSSIBILITY inferences they are predicted to give rise to on the standard implicature approach.

Specifically, on the first parse (without $\mathcal{O}$), the predictions were exactly the same as before; in particular, one of the UNCERTAINTY inferences, namely $\neg\Box_s A$, was false in both TARGET conditions, whereas one of the POSSIBILITY inferences, namely $\Diamond_s B$, was false only in the TARGET-2 conditions. By contrast, on the second parse (with $\mathcal{O}$), the corresponding inferences, namely $\neg\Box_s A$ and $\Diamond_s \mathcal{O}(B)$, were both false in both TARGET conditions; in particular, the POSSIBILITY inference $\Diamond_s \mathcal{O}(B)$ was false in the TARGET-1 conditions because, in these conditions, the speaker was always certain that A was true (i.e., $B \wedge \neg A$ is impossible). Thus, we reasoned that if participants generally opt for the first parse, the key contrasts observed in Experiments 1 and 2 should reproduce in this experiment. On the other hand, if participants opt instead for the second parse, these contrasts should disappear, or at least shrink in a noticeable way. The first case would support our current interpretation of the results, whereas the second would challenge it in favor of the alternative explanation discussed at the end of the previous section.

(11) a. **Current interpretation**
Participants opt for the first parse. The contrasts from Exps. 1-2 obtain because POSSIBILITY may arise independently of CERTAINTY.
Prediction: similar contrasts should reproduce in Exp. 3

b. **Alternative interpretation**
The contrasts from Exps. 1-2 are consistent with POSSIBILITY living on CERTAINTY. They obtain because participants opt for the second parse.
Prediction: no such contrasts should be found in Exp. 3

Before proceeding, let us briefly discuss two alternative parses for our test sentences, which can be set aside here due to their low plausibility. First, the meaning of each individual disjunct could be strengthened via a covert instance of ‘only’, as shown in (12). The reading associated with this parse amounts to an exclusive reading of the disjunction, which in our experiments would be false in the TARGET-1 conditions and true in the TARGET-2 conditions. Thus, should participants opt for this strategy, we would see a contrast between the TARGET-1 and TARGET-2 conditions in the opposite direction to the one we currently expect and previously observed. The results so far and those to come are incompatible with participants opting for this parse.

(12) a. The mystery box contains ⟨only⟩ a blue ball or ⟨only⟩ a yellow ball.

b. $\mathcal{O}(B) \vee \mathcal{O}(A) \equiv (B \wedge \neg A) \vee (A \wedge \neg B)$

Second, in place of the first disjunct, the meaning of the second disjunct could be strengthened, as shown in (13). If this sort of parse were available, it would be problematic for our present purposes. In particular, for the same reasons as those discussed for $\mathcal{O}(A) \vee B$, the availability of this parse would cast doubt on the ability of our novel TARGET-1 conditions to discriminate between UNCERTAINTY and POSSIBILITY inferences, as all of them would be true in these conditions.

(13) a. The mystery box contains a blue ball or ⟨only⟩ a yellow ball.

b. $B \vee \mathcal{O}(A) \equiv B \vee (A \wedge \neg B)$

However, evidence from Singh (2008) suggests that, unlike its symmetric variant, a parse like (13) is not readily available, even in cases where it would prevent infelicity from arising. To illustrate this point, consider the minimal pair in (14):[11]

(14) a. The mystery box contains a yellow ball, or both a yellow ball and a blue ball.
b. #The mystery box contains both a yellow ball and a blue ball, or a yellow ball.

These sentences are examples of so-called "Hurford" disjunctions: in both cases, one of the disjuncts entails the other, at least superficially. Crucially, as Singh (2008) observes, a sentence like (14-a), of the form 'A or (A and B)', is fully felicitous, whereas its variant in (14-b), of the form '(A and B) or A', isn't. This asymmetry is accounted for if we assume that (14-a) genuinely allows for a parse that breaks the relation of entailment between the two disjuncts, namely $\mathcal{O}(A) \vee (A \wedge B)$, whereas (14-b) doesn't, i.e., $(A \wedge B) \vee \mathcal{O}(A)$ is not available or at least not to the same extent. We take these observations to suggest that, while an 'only'-reading of the first disjunct is possible for our test sentences, an 'only'-reading of the second disjunct of the sort illustrated in (13) is very unlikely.

6.1 Participants

100 novel participants took part in this study (mean age 40 yrs; 49 female). Participants were recruited online through Prolific using the same prescreening criteria as in Experiments 1 and 2. Participants were paid £1.20, and median completion time was about 8 minutes (hourly rate: £9/hr). The consent and data collection procedures were the same as in Experiments 1 and 2.

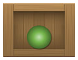
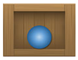
6.2 Materials and design

The materials from Experiment 2 were modified in two ways. First of all, as explained above, we reversed the order of the disjuncts in the test sentences, changing it from $A \vee B$ to $B \vee A$. These sentences were associated with the same picture types and subtypes as in Experiment 2. As a result, Table 5 also stands as an illustrative summary of the test conditions in Experiment 3, provided the modification of the test sentences that we just described. Concretely, the example sentence in Table 5, *'The mystery box contains a yellow ball or a blue ball*, should now read as *'The mystery box contains a blue ball or a yellow ball'*.

Second, we modified the pictures used in the BAD conditions for the $A \wedge B$ controls (C3) to further control for low-level response strategies of the sort previously discussed. Specifically, the relevant pictures were modified so as to depict A-balls and B-balls only, as shown in Table 8. We reasoned that, if a participant primarily focuses their attention on whether the colors of the balls depicted on the picture match the color terms used in the sentence, they should incorrectly accept these cases. While this color-matching strategy appears unlikely given the previous results to the

[11] We thank one of our anonymous reviewers for pointing this out to us.

Table 8 Schematic description and illustration of the pictures types used in the BAD conditions for the C3 controls in Exps. 1-2 and in Exp. 3

Sentence	Condition	Experiment	Example picture			
C3: $A \wedge B$	BAD	Exp. 1-2	 A	 CD	 C	 ?
		Exp. 3	 A	 AA	 B	 ?

Sentence: *'The mystery box contains a yellow ball and a blue ball.'*

C4-GOOD controls, the use of negation in the C4 sentences introduces an extra layer of complexity which may prevent participants from adopting this response strategy. These novel cases allowed us to probe again for this strategy in trials involving positive sentences structurally similar to our test sentences.

The rest of the design was identical to that of Experiments 1 and 2 in all respects. As in the previous experiments, each control condition was instantiated 3 times, resulting in 24 control trials. The TARGET-1, TARGET-2*, and FALSE* conditions were instantiated 6 times each, while the TRUE* and TRUE-EXCL conditions were instantiated 3 times each, resulting in 24 test trials. Instances of the 2 target conditions were evenly distributed across their respective sub-conditions. Thus, the experiment included a total of 48 trials.

6.3 Procedure

The procedure was identical to the one used in Experiments 1 and 2 (see Sect. 4.3 for details).

6.4 Data availability

The stimuli, data and analysis code associated with Experiment 3 are available open access on the OSF platform at https://osf.io/4ut2c/.

6.5 Data treatment and analysis

The data was treated and analyzed using the data analysis pipelines from Experiments 1 and 2. Responses from 18 participants were removed because their performance in the control trials did not reach the pre-established threshold of 80% accuracy. The performance of the remaining subjects (n=82) was high both in the BAD conditions (M=92.8%, 95% CI[91.0, 94.2]) and the GOOD conditions (M=97.7%, 95% CI[96.5, 98.4]). We note, however, that while there was little variation among the GOOD controls (all Ms>95%), the mean accuracy rate was slightly lower for the C3-BAD controls (M=84.1%, 95% CI[79.1, 88.2]) than the other BAD controls (all Ms>92%).

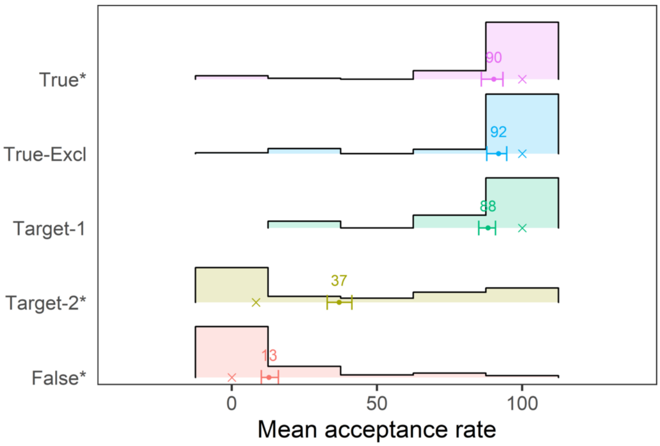


Fig. 4 Mean acceptance rate (i.e., proportion of 'Good' responses) to target sentences by experimental condition in Experiment 3. For each condition, the distribution of by-participant mean rates is visualized by a histogram, the grand mean by a dot with its rounded value on top and the 95% CI around it, and the median by a cross

As explained above, a participant performing poorly on the C3-BAD controls is suggestive that they relied on low-level response strategies, at least on occasion. For this reason, participants' responses in the test trials were also analyzed by including only those participants who performed perfectly on all C3 control trials ($n = 63$). This analysis is reported in the analysis script associated with Experiment 3. The results show that the mean acceptance rates per condition were similar to those observed in the main analysis that we report on below. We take these results to rule out any explanation of the following results in terms of color-matching strategy.

6.6 Results

Figure 4 shows the mean acceptance rate (i.e., the proportion of 'Good' responses) for the test sentences as a function of the experimental condition. The outputs of the statistical models and analyses are summarized in Table 9.

The results were similar to those from Experiment 2. The mean acceptance rate in the TARGET-1 conditions was very high (M=88.21%, 95% CI[85.1, 90.8]), similar to those found in the TRUE-EXCL conditions (M=91.9%, 95% CI[87.8, 94.7]) and the TRUE* conditions (M=90.2%, 95% CI[85.9, 93.4]), with no significant difference between TRUE* and TRUE-EXCL ($z = -0.10$, $p = 0.92$).[12] By contrast, the mean acceptance rate in the TARGET-2* conditions was in the lower mid-range (M=37.0%, 95% CI[32.8, 41.3]), falling between the highest rates observed in the TARGET-1, TRUE-EXCL and TRUE* conditions and the lowest rate observed in the FALSE* conditions (M=12.8%, 95% CI[10.1, 16.0]). In particular, we found no significant difference between TARGET-2* and TARGET-1 ($z = -6.54$, $p < 0.001$), and the difference

[12] As in Experiments 1 and 2, the by-participant mean rates in the TARGET-2(*) conditions were not distributed unimodally ($D = 0.12$, $p < .001$), as shown in Fig. 4 by the two peaks in the distribution.

Table 9 Summary of statistical models outputs and analyses. The R pseudo-code for the `glmer` model is `Response` $\sim$ 1 +`Condition`+ (1 + `Condition|Subject`). The coding scheme is successive difference coding on `Condition`. z-values are calculated based on the estimated coefficients and their standard errors. p-values are calculated using Wald tests. Confidence intervals are calculated using the Wald method. For random effects estimates, refer to the analysis script. Conditional and Marginal R^2 values are 0.91 and 0.5, respectively

	Estimate	95% CI	z-value	p-value
Intercept	1.06	[0.07, 2.05]	2.10	0.036
TRUE*.TRUE-EXCL	−0.13	[−2.51, 2.26]	−0.10	0.92
TRUE-EXCL.TARGET-1	−0.62	[−2.23, 1.00]	−0.76	0.46
TARGET-1.TARGET-2*	−6.45	[−8.39, −4.52]	−6.54	**<.001**
TARGET-2*.FALSE*	−3.24	[−5.87, −0.61]	−2.41	0.016

between TARGET-2 and FALSE* was weakly significant ($z = -2.41$, $p = 0.016$). In sum, these results reproduce the main key contrasts observed in the previous experiments.

6.7 Discussion

Experiment 3 aimed to control for another possible parse of the test sentences on which the first disjunct receives an 'only'-interpretation. For these purposes, we modified the materials used in Experiment 2 by swapping the position of the disjuncts in the test sentences, changing it from 'A or B' to 'B or A', while keeping the sentence-picture combinations the same as before. As we explained, this minimal change in design made it possible to discriminate between the two parses of interest, $B \vee A$ and $\mathcal{O}(B) \vee A$, and distinguish the predictions of two competing interpretations for the contrasts observed in Experiments 1-2 (see (11) for a summary).

The results that we obtained were essentially similar to those of Experiments 1-2, showing that reversing the order of disjuncts did not affect participants' responses in any significant way. In particular, participants consistently accepted the novel 'B or A' sentences in the TARGET-1 conditions, while they robustly rejected them in the TARGET-2* conditions. These findings are incompatible with participants opting for the alternative parse that we discussed and, in turn, undermine any explanation of the contrasts from Experiments 1-2 based on this parsing strategy. These findings, on the other hand, align with our initial interpretation of the data and suggest that the main conclusion from the previous experiments should be maintained: disjunctive sentences can give rise to POSSIBILITY in the absence of UNCERTAINTY.

7 General discussion

The experiments discussed in Sects. 4, 5 and 6 pose a clear challenge to the traditional implicature approach: simple disjunctive sentences can give rise to POSSIBILITY inferences, even in cases where UNCERTAINTY inferences are not derived. For the sake of clarity, we repeat below an example with the relevant inferences.

(15) The mystery box contains a yellow or a blue ball. $A \vee B$

a. $\neg\Box_s A \wedge \neg\Box_s B$ UNCERTAINTY

b. $\Diamond_s A \wedge \Diamond_s B$ POSSIBILITY

To begin our discussion, we first draw a connection between the findings presented in this work and similar observations that have been discussed in relation to disjunction in embedded contexts.

7.1 The connection with distributive inferences

In this work, we focused on cases on plain disjunction and its related inferences. However, embedded disjunction displays similar patterns. We illustrate this by means of a disjunction embedded under a universal modal.[13]

(16) The mystery box must contain a blue ball or a yellow ball. $\Box(A \vee B)$

a. $\neg\Box A \wedge \neg\Box B$ NEGATED UNIVERSAL

b. $\Diamond A \wedge \Diamond B$ DISTRIBUTIVE

The example in (16) gives rise to the NEGATED UNIVERSAL inference in (16-a), corresponding to the UNCERTAINTY inferences discussed here for plain disjunction and illustrated in (15-a). Likewise, (16) is also associated with the existential inference in (16-b), often called DISTRIBUTIVE inference, corresponding to our POSSIBILITY inferences in (15-b).

In the case of embedded disjunction, the traditional implicature approach predicts that the lower bound DISTRIBUTIVE inferences are dependent on the upper bound NEGATED UNIVERSAL part, similarly to the predictions for plain disjunction. However, experimental research (Crnič et al. 2015; Marty et al. 2023) found that DISTRIBUTIVE inferences can be derived even in the absence of NEGATED UNIVERSAL ones.

In light of the generality of the phenomena, we would want a theory that comprehensively captures the independent derivation of lower bound distributive/possibility inferences in a unified fashion. We start by looking at recent implicature accounts (Crnič et al. 2015; Bar-Lev and Fox 2023) that have been designed to predict DISTRIBUTIVE inferences in the absence of NEGATED UNIVERSAL ones. We then consider whether these accounts can be extended to IGNORANCE inferences. Subsequently, we shift our focus to non-implicature accounts, which do not derive POSSIBILITY as an implicature. To make the discussion concrete, we focus on a recent proposal by Aloni (2022).

---

[13] Note that distributive inferences can also arise in the case of nominal quantifiers, as in (i):

(i) Every visible box contains a blue ball or a yellow ball. $\forall x(P(x) \vee Q(x))$

a. $\neg\forall x P(x) \wedge \neg\forall x Q(x)$ NEGATED UNIVERSAL

b. $\exists x P(x) \wedge \exists x Q(x)$ DISTRIBUTIVE

Crnič et al. (2015) showed that DISTRIBUTIVE inferences do indeed arise for nominal quantifiers in the absence of the NEGATED UNIVERSAL. They also argued that DISTRIBUTIVE inferences do not arise in the case of modals. Marty et al. (2023) replicated the results of Crnič et al. (2015) for the nominal case, but they also showed that distributive inferences do indeed arise for the modal case, as discussed at the beginning of this section.

7.2 Implicature accounts

7.2.1 Distributive inferences and existential alternatives

In this section, we outline how a recent account by Bar-Lev and Fox (2023) in the grammatical approach can derive DISTRIBUTIVE inferences without the need of deriving NEGATED UNIVERSAL ones. The core of the proposal in Bar-Lev and Fox (2023) is the application of recursive exhaustification, building on previous work by Bar-Lev and Fox (2020) and Fox (2007). In this approach, implicatures are computed by applying a covert exhaustivity operator, labeled EXH,[14] which negates innocently excludable (IE) alternatives (i.e., those alternatives to ϕ that can be negated simultaneously without contradicting ϕ or entailing the other alternatives):

(17) a. $IE(\phi, S) := \bigcap \left\{ S' \;\middle|\; \begin{array}{l} S' \subseteq S \text{ and } S' \text{ is a maximal subset of } S \\ \text{such that } \{\neg\psi : \psi \in S\} \cup \{\phi\} \text{ is consistent} \end{array} \right\}$

b. $[\![\text{EXH}(\phi)]\!](w) = [\![\phi]\!](w) \wedge \forall\psi \in IE(\phi, ALT(\phi)) : \neg[\![\psi]\!](w)$

In the following, we illustrate the derivation of DISTRIBUTIVE inferences for the modal case in example (16). The alternatives associated with $\Box(A \vee B)$ are the full set of alternatives in (18). Note that the existential alternatives in blue ($\Diamond A$ and $\Diamond B$) are essential for the derivation of POSSIBILITY.

(18) $Alt(\Box(A \vee B)) = \{\Box A, \Box B, \Diamond A, \Diamond B, \Box(A \wedge B), \Box(A \vee B), \Diamond(A \vee B), \Diamond(A \wedge B)\}$

According to Bar-Lev and Fox (2023), the derivation of DISTRIBUTIVE inferences is achieved by using a double application of the exhaustivity operator EXH, along with the relevant pruning of alternatives, which we discuss below.

In particular, the conjunctive alternative $\Diamond(A \vee B)$ in (18), in red, needs to be pruned to avoid the derivation of EXCLUSIVITY, which, as we discuss in more detail in Sect. 7.2.4, would also derive NEGATED UNIVERSAL in combination with DISTRIBUTIVE.[15] As a result, the pruned set of alternatives, on which EXH operates, looks as in (19):[16]

[14] The EXH operator can be viewed as a more elaborate version of the simple $\mathcal{O}$ operator discussed at the end of Sect. 5.

[15] The pruning of these alternatives is allowed, as it does not break symmetry. Two alternative ψ_1 and ψ_2 of ϕ are symmetric when $\psi_1 \vee \psi_2 \equiv \phi$ and $\psi_1 \wedge \psi_2 \equiv \bot$.

An alternative explanation could link the pruning mechanism to the questions under discussion, instead of/in addition to symmetry breaking. For instance, it might be possible that conjunctive alternatives of the form $A \wedge B$ were not relevant to the QUD, but the individual alternatives A and B were.

[16] To also avoid weaker exclusivity inferences of the form $\neg\Box(A \wedge B)$, the alternative $\Box(A \wedge B)$ should also be pruned. While this was not tested in our experiments, the absence of such an inference has been observed in previous work by Crnič et al. (2015), which we expect to align with the current experimental paradigm. One reviewer appropriately points out that the set of alternatives with both conjunctive alternatives being pruned is not problematic for innocent exclusion, which we are assuming here, but it would be problematic if we also assume innocent inclusion (Bar-Lev and Fox 2020). Since no alternative would be excludable and the conjunction of all alternatives is consistent, EXH with innocent inclusion would give us the conjunction of all alternatives, and thus $\Box A$ and $\Box B$ would end up being entailed. Innocent exclusion allows for deriving DISTRIBUTIVE without any form of exclusivity by pruning both conjunctive

(19) $Alt_{pruned}(\Box(A \vee B)) = \{\Box A, \Box B, \Diamond A, \Diamond B, \Box(A \vee B), \Diamond(A \vee B), \Box(A \wedge B)\}$

Distributive inferences are then obtained by recursive exhaustification, as in (20). On the first application of EXH, only the $\Box(A \wedge B)$ alternative is excludable. With the subsequent application of EXH, thanks to the presence of the individual existential alternatives, DISTRIBUTIVE inferences are then generated:

(20) $\text{EXH}(\text{EXH}\Box(A \vee B))$
$\equiv \Box(A \vee B) \wedge \neg\text{EXH}(\Box A) \wedge \neg\text{EXH}(\Box B) \wedge \neg\Box(A \wedge B)$
$\equiv \Box(A \vee B) \wedge \neg(\Box A \wedge \neg\Box B \wedge \neg\Diamond B) \wedge \neg(\Box B \wedge \neg\Box A \wedge \neg\Diamond A) \wedge \neg\Box(A \wedge B)$
$\equiv \Box(A \vee B) \wedge \Diamond A \wedge \Diamond B \wedge \neg\Box(A \wedge B)$
$\Rightarrow \Box(A \vee B) \wedge \Diamond A \wedge \Diamond B$

We have seen how an implicature approach in the grammatical tradition can account for the independent derivation of DISTRIBUTIVE inferences. What remains to determine is how this account can be extended to the case of plain disjunctive sentences and POSSIBILITY inferences discussed in our experimental studies. Before moving to that, we discuss another recent variant of the implicature approach by Crnič et al. (2015).

7.2.2 Distributive inferences and embedded EXH

We now outline a different implicature approach (Crnič et al. 2015) that does not rely on the presence of existential alternatives,[17] but on the possibility of embedding the EXH operator, as in (21):

(21) a. The mystery box must contain A or B.
b. $\text{EXH}_1(\Box(\text{EXH}_2(A \vee B)))$

In (21-b), we introduce two indices for the EXH operator to distinguish between the set of relevant alternatives on which the two EXH operate. In particular, the set of alternatives relevant for EXH_1 is the following:

(22) $Alt_{\text{EXH}_1}(\Box(\text{EXH}_2(A \vee B))) = \{\Box(\text{EXH}_2(A)), \Box(\text{EXH}_2(B)), \Box(\text{EXH}_2(A \wedge B))\}$
a. $\Box(\text{EXH}_2(A)) \equiv \Box(A \wedge \neg B)$
b. $\Box(\text{EXH}_2(B)) \equiv \Box(B \wedge \neg A)$
c. $\Box(\text{EXH}_2(A \wedge B)) \equiv \Box(A \wedge B)$

The crucial assumption to obtain the equivalences in (22) is that EXH_2 operates on the restricted set of alternatives in (23) where the conjunctive alternative is absent:

(23) $Alt_{\text{EXH}_2}(\psi) = \{A, B\}$

alternatives. To obtain the same result, innocent inclusion needs also to prune the $\Box A$ and $\Box B$ alternatives, having only $\Diamond A$ and $\Diamond B$ in the set of alternatives.

[17] We refer to Bar-Lev and Fox (2023) for criticism of Crnič et al. (2015) and the independent need of existential alternatives.

Given these assumptions, we can obtain POSSIBILITY as desired:

(24) $\text{EXH}_1(\Box(\text{EXH}_2(A \vee B)))$
$\equiv \Box(A \vee B) \wedge \neg\Box(A \wedge \neg B) \wedge \neg\Box(B \wedge \neg A) \wedge \neg\Box(A \wedge B)$
$\equiv \Box(A \vee B) \wedge \Diamond(\neg A \vee B) \wedge \Diamond(\neg B \vee A) \wedge \neg\Box(A \wedge B)$
$\equiv \Box(A \vee B) \wedge \Diamond A \wedge \Diamond B \wedge \neg\Box(A \wedge B)$
$\Rightarrow \Box(A \vee B) \wedge \Diamond A \wedge \Diamond B$

Similarly to the recursive EXH approach, pruning the conjunctive alternative from (22) would remove the weaker exclusivity inference derived above.

7.2.3 Extending the implicature approach to IGNORANCE

A natural way to extend the implicature approach to plain disjunctive sentences is the use of a silent doxastic operator (Meyer 2013; Buccola and Haida 2019, among others), which we indicate with $\Box_s$. The fundamental assumption is that every assertively used sentence is associated with a covert doxastic operator $\Box_s$, where the subscript refers to the doxastic source (in this case s stands for the speaker). For instance, an assertion like 'It is raining' will be rendered as $\Box_s p$, meaning that in all the doxastic possibilities of the speaker, it is the case that it is raining.

We first discuss the results and derivation for the recursive approach outlined in Sect. 7.2.1, and we then consider the embedded EXH account discussed in Sect. 7.2.2.

Under the extension of the first account, a plain disjunctive sentence is associated with the alternatives in (25), where $\Diamond_s$ corresponds to the existential counterpart of $\Box_s$. As discussed for the case of distributive inferences, we have pruned the alternative responsible for EXCLUSIVITY.

(25) $Alt_{pruned}(\Box_s(A \vee B)) = \{\Box_s A, \Box_s B, \Diamond_s A, \Diamond_s B, \Box_s(A \vee B), \Diamond_s(A \vee B), \Box_s(A \wedge B)\}$

By virtue of this silent operator, the derivation of the POSSIBILITY inferences $\Diamond_s A$ and $\Diamond_s B$ parallels the case discussed in (20), thus capturing our results.

One main question for the extension of the account by Bar-Lev and Fox (2023) to IGNORANCE inferences concerns the nature of the silent operator involved in the alternatives. While it is not particularly controversial to assume a silent belief operator $\Box_s$ for assertively used sentences, the use of $\Diamond_s$ as an implicit or covert operator that scopes over sentences is considerably less established. One way to make sense of $\Diamond_s$ is as a weak form of assertion (Incurvati and Schlöder 2019; Dorst and Mandelkern 2022; Mandelkern and Dorst 2022), the idea that language users can (weakly) assert something even when they have low rational credence in it. However, assuming that $\Diamond_s$ is also silent, the existence of such an operator would allow us to utter a sentence like (26) also when it is only compatible with our belief that it is raining ($\Diamond_s p$), which is arguably not the case.

(26) It is raining. $\Diamond_s / \Box_s p$

The use of the existential alternative $\Diamond_s$ thus poses a challenge for the recursive approach.[18]

The embedded approach discussed in Sect. 7.2.2 fares better in this regard since it does not rely on existential alternatives. Analogously to the case with an overt modal discussed in Sect. 7.2.2, a plain disjunctive sentence like (27-a) is associated with two EXH operators: one applied before the covert doxastic operator $\Box_s$, and one applied after.

(27) a. The mystery box contains A or B.
b. $\text{EXH}_1(\Box_s(\text{EXH}_2(A \vee B)))$

The derivation is then parallel to the one outlined in Sect. 7.2.2, with the crucial assumption of pruning the conjunctive alternative from the domain of the embedded EXH_2.[19]

In conclusion, the extension of the implicature approach to IGNORANCE relies on a silent doxastic operator $\Box_s$ and leads to significant theoretical consequences. On the one hand, the recursive EXH approach would lead us to posit a weak covert doxastic operator $\Diamond_s$. On the other hand, the embedded EXH approach would lead us to require EXH operators to be adjoined before and after a silent operator, which has been previously considered in the literature (Meyer 2013).

Before moving to the next section, we would like to mention a possible alternative proposal that does not rely on pruning.[20] Instead of pruning the alternatives, we might devise a mechanism of implicature computation that is sensitive to the question under discussion (QUD). Specifically, if we assume that the QUD of the experiment is of the form ‘What *might* be inside the box?’ rather than ‘What *is* inside the box?’, participants might only generate the weakest inferences of a disjunction ‘A or B’ that are sufficient to solve the issue raised by the QUD.

When considering control sentences that do not contain a disjunction, particularly the C1-BAD controls (refer to Table 3), we observe that, as discussed in fn. 5, these controls are true under a ‘might-QUD’, and the high acceptance rates by some participants may be indicative of this interpretation. In our analysis, we adopted a conservative approach and excluded such participants. Assuming that the remaining participants did not engage in might-QUD reasoning, we showed that the primary effects of interest were present. Nonetheless, it remains possible that participants responded accurately to these controls for reasons unrelated to the QUD or that different QUDs may apply to different types of sentences.

A possible way to situate this idea is the QUD-sensitive account discussed by Cremers et al. (2022) and Cremers et al. (2023). We believe this is a step in the right direction, both empirically, by designing experimental frameworks that control for

[18] One possibility to explore for this approach in relation to this challenge is the observation that strong assertions are often unmarked, whereas weak assertions require an overt marker like ‘perhaps’, as in the sentence ‘Perhaps it is raining’ ($\Diamond_s p$), which can indeed be uttered when it is only compatible with our belief that it is raining.

[19] Importantly, including the conjunctive alternative in the embedded exhaustification operator would yield UNCERTAINTY, as well as EXCLUSIVITY (see, e.g., Meyer 2013, pp. 52–54).

[20] We would like to thank Roni Katzir and one anonymous reviewer for pointing us in this direction.

the QUD, and theoretically, by developing QUD-sensitive theories. A proper implementation of this proposal must be left for a future occasion.

7.2.4 The role of EXCLUSIVITY

In the previous discussion, we pruned the alternatives to prevent the generation of EXCLUSIVITY. The reason behind this decision was that EXCLUSIVITY, when combined with POSSIBILITY or DISTRIBUTIVE inferences, leads to UNCERTAINTY or NEGATED UNIVERSAL inferences, as in (28). This would have been undesirable because we aim to generate POSSIBILITY without also generating UNCERTAINTY.

(28) The role of EXCLUSIVITY

a. $\Diamond A \wedge \Diamond B$

b. $\Box\neg(A \wedge B)$

c. $\rightsquigarrow \neg\Box A \wedge \neg\Box B$

We illustrate this for the first approach illustrated above, which assumes existential alternatives. The derivation with the full set of alternatives in (18) is given in (29). In this case, on the first application of EXH, $\Diamond(A \wedge B)$ and $\Box(A \wedge B)$ are excludable, generating EXCLUSIVITY. As before, the subsequent application of EXH yields distributive inferences:

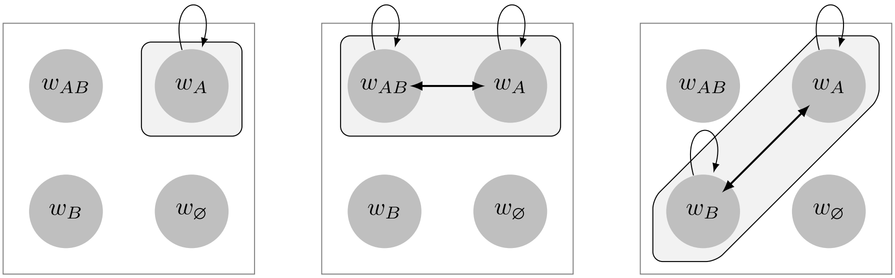
(29) $\text{EXH}(\text{EXH}\Box(A \vee B))$
$= \Box(A \vee B) \wedge \neg\Box(A \wedge B) \wedge \neg\Diamond(A \wedge B) \wedge \neg\text{EXH}(\Box A) \wedge \neg\text{EXH}(\Box B)$
$= \Box(A \vee B) \wedge \neg\Box(A \wedge B) \wedge \neg\Diamond(A \wedge B) \wedge \neg(\Box A \wedge \neg\Box B \wedge \neg\Diamond B) \wedge \neg(\Box B \wedge \neg\Box A \wedge \neg\Diamond A)$
$= \Box(A \vee B) \wedge \neg\Diamond(A \wedge B) \wedge \Diamond A \wedge \Diamond B$

As such, this approach predicts that UNCERTAINTY follows from EXCLUSIVITY (i.e., UNCERTAINTY cannot arise without the presence of EXCLUSIVITY). As regards our experiments, we found no evidence of either UNCERTAINTY or EXCLUSIVITY.[21] An important question, which we explore in Sect. 7.5, is whether there can be cases in which UNCERTAINTY is derived independently of EXCLUSIVITY.

It is worth mentioning that the mechanism behind the derivation of both EXCLUSIVITY and POSSIBILITY is the same, in contrast to the non-implicature accounts presented in Sect. 7.3. In those cases, the derivation of EXCLUSIVITY is regarded as a type of implicature and is generally separated from distributive inferences.

We now outline a recent non-implicature approach (Aloni 2022) to IGNORANCE inferences that accounts independently for POSSIBILITY and related inferences. We also explore how such a system can be enriched with a mechanism of implicature computation that is able to derive UNCERTAINTY via EXCLUSIVITY. It is worth noting that while our experimental data seem to suggest an independent mechanism to generate POSSIBILITY, the choice to take Aloni's 2022 account as an example of a non-implicature approach is not crucial, and similar considerations can be explored using alternative accounts, such as the one discussed in Goldstein (2019).

[21] As discussed in fn. 16, we found no evidence of exclusivity inferences $\Box_s\neg(A \wedge B)$. In previous experiments on distributive inferences, Crnič et al. (2015) also found no evidence of weak exclusivity inferences $\neg\Box_s(A \wedge B)$, meaning that the configuration AB-AB-AB-? was judged as good.



(a) $s_1 \vDash A \vee B s_1 \nvDash (A \vee B)^+$ (b) $s_2 \vDash A \vee B s_2 \vDash (A \vee B)^+$ (c) $s_3 \vDash A \vee B s_3 \vDash (A \vee B)^+$

Fig. 5 Illustrations. A state s is a set of possible worlds (in w_{AB} both A and B hold, in w_A only A holds and so on). The accessibility relation for each world is indicated with an arrow.

7.3 A non-implicature account

On Aloni's 2022 account, POSSIBILITY and related inferences do not arise from Gricean reasoning or the application of a covert exhaustification operator. Instead, these inferences are the result of another factor referred to as neglect-zero. In particular, Aloni (2022) puts forward the idea that language users, when interpreting a sentence, construct structures that represent reality, and they thus systematically disregard structures that vacuously satisfy the sentence due to an empty configuration, also referred to as zero-models. Neglect-zero relates to a tendency that has been observed in number cognition and language development with regard to the number zero, as discussed in Nieder (2016).

These structures are built upon the notion of a state, a set of possible worlds, which reflects the speaker's information state and upon which formulas are interpreted. For instance, Fig. 5 depicts three states: $s_1 = \{w_A\}$, $s_2 = \{w_{AB}, w_A\}$ and $s_3 = \{w_A, w_B\}$. As said, states are meant to encode the epistemic state of the speaker. For instance, in s_3, the speaker considers possible that only A holds or that only B holds, while in s_2 and s_1, the speaker is certain that A holds. Note that this notion of states provides a natural correspondence to the information state of the child's character in our mystery box experiments. The three states in Fig. 5 could correspond to the box configurations '$A - A - A-?$', '$A - AB - A-?$', and '$A - B - B-?$', respectively.

Aloni (2022) formalizes this in a Bilateral State-based Modal Logic, which we refer as BSML. We focus on the fundamental tenets relevant to us, and refer the reader to Aloni (2022) for further details.

Atomic formulas are true in a state when all worlds in the state make the formula true. For instance, in Fig. 5, A holds in s_1 and s_2, but not in s_3, since A is not true in w_B. BSML employs a split notion of disjunction, which is supported in a state s when the state can be split into two sub-states (i.e., s is the union of two sub-states), each supporting one of the disjuncts. All the states in Fig. 5 support a disjunction of the form $A \vee B$. For the state s_1 in Fig. 5(a), the two sub-states would be $\{w_A\}$ supporting A and the empty state $\emptyset$ supporting B, since it vacously holds that all the worlds in the state support B.[22]

[22]Note that $\{w_A\} \cup \emptyset = \{w_A\} = s_1$.

Lastly, we note that each world can see/access other worlds according to the accessibility relation specified in the model, as represented in Fig. 5 by means of an arrow. Roughly, a universal modal $\Box A$ is true when all worlds in the state see A-worlds. A possibility modal $\Diamond A$ is true when all worlds in the state see some A-world.

7.3.1 Plain disjunction and possibility

As said above, Aloni (2022) proposes that language users neglect zero-models and therefore rule out the possibility of verifying a sentence by means of empty configurations. Aloni (2022) implements this neglect-zero effect by means of an enrichment function $(\cdot)^+$ which excludes the empty set as a possible verifier. Formally, this can be defined recursively on the complexity of the formulas, but here we focus on the case of disjunction.

An enriched disjunction $(A \vee B)^+$ is supported in a state when the state can be split into two *non-empty* sub-states, each supporting one of the disjuncts. For instance, the disjunction $A \vee B$ is no longer supported in s_1 in Fig. 5 when enriched (i.e., $(A \vee B)^+$), since as we have discussed above one of the supporting sub-states is the empty set.

As a result, an enriched disjunction $(A \vee B)^+$ entails POSSIBILITY, but not UNCERTAINTY, where the entailment $\phi \models \psi$ in (30) roughly means that if a state in a model makes ϕ true, it also makes ψ true:

(30) a. $(A \vee B)^+ \models \Diamond_s A \wedge \Diamond_s B$
b. $(A \vee B)^+ \not\models \neg\Box_s A \wedge \neg\Box_s B$

(30-a) can be formally proved in BSML when the accessibility relation is state-based, meaning that each world in the state sees all the other worlds in the state. State-basedness aligns with the view that speakers are aware of their information state, and allows us to capture the fact that A is epistemically possible when A is true in some world in the speaker's information state. Assuming that the accessibility relation is state-based and s stands for the speaker, we use $\Diamond_s A$ to indicate that A is possible in the information state of the speaker, in line with our previous notation of epistemic modality.

A pragmatically enriched disjunction is supported only in states that include both A-worlds and B-worlds, and this ensures that $\Diamond_s A$ and $\Diamond_s B$ hold. However, for (30-b), a pragmatically enriched disjunction is also supported in states like Fig. 5(b), which do not support $\neg\Box_s A$, since A is true in all words in the state.

Therefore, BSML offers a systematic explanation for the experimental findings presented in Sects. 4, 5 and 6, namely that POSSIBILITY inferences are derived as neglect-zero effects without UNCERTAINTY inferences.

7.3.2 Embedded disjunction and distributive inferences

In Sect. 7.1, we observed that there is a strong parallelism between inferences related to plain disjunction and those in an embedded context. Therefore, a uniform theory that captures both types of inferences is desirable. BSML readily accounts for DISTRIBUTIVE inferences, since $[\Box_{(s)}(A \vee B)]^+$ entails $\Diamond_{(s)} A$ and $\Diamond_{(s)} B$ both for epistemic and other modalities.[23] The reason is that when $[\Box_{(s)}(A \vee B)]^+$ holds, all worlds in the state must access both A-worlds and B-worlds, due to the pragmatic enrichment affecting the disjunction, thus making $\Diamond_{(s)} A$ and $\Diamond_{(s)} B$ also satisfied (Aloni 2022):[24]

(31) a. $[\Box(A \vee B)]^+ \models \Diamond A \wedge \Diamond B$
b. $[\Box(A \vee B)]^+ \not\models \neg\Box A \wedge \neg\Box B$

Therefore BSML can provide a unified account of ignorance and distributive inferences.

7.3.3 The role of exclusivity

The non-implicature framework explored so far accounted for the derivation of POSSIBILITY and related inferences in the absence of the corresponding negated upper bound inferences and EXCLUSIVITY. While this approach makes the correct predictions for our experimental paradigm, there are instances in which EXCLUSIVITY and UNCERTAINTY inferences are derived. We return to this issue in Sect. 7.5.

For the moment, we observe that adding EXCLUSIVITY to either POSSIBILITY or DISTRIBUTIVE inferences leads to the derivation of UNCERTAINTY or NEGATED UNIVERSAL inferences. This aligns with our earlier observations in Sect. 7.2.4 within the context of the grammatical approach. Aloni (2022) does not provide an account of scalar implicatures. An important next step is to extend non-implicature accounts by incorporating a scalarity operator $(\cdot)^\sigma$, which enables the generation of scalar implicatures (such as EXCLUSIVITY for a plain disjunction). In this way, the interplay between the neglect-zero enrichment $(\cdot)^+$ and the scalar enrichment $(\cdot)^\sigma$ could be adequately studied.

In the subsequent section, we delve into an alternative account in the context of game-theoretic approaches to pragmatics, which builds upon Grice's original insights regarding cooperation between the speaker and hearer. This theory offers an additional viewpoint on the behavior of disjunction and its associated inferences.

7.4 The speaker-hearer perspective and linking hypotheses

A pivotal aspect of experimental research lies in identifying the appropriate linking hypotheses for an experimental paradigm. These hypotheses establish the connection between the measured response and the theoretical frameworks that we aim to enhance our understanding of. In this section, we revisit our experimental design to explore why certain inferences are derived, while others are not, and examine how recent game-theoretic approaches may enhance our understanding of these patterns.

In the current experimental paradigm, participants consider what is being said and conveyed by the child character's utterance and then assess whether or not the result aligns with what can be assumed about the character's information state.

[23] In this case, we are using $\Box_s$ for an overt epistemic modal and $\Box$ for an overt deontic modal.

[24] Note also that first-order extensions of Aloni's 2022 system, such as Aloni and van Ormondt (2023), can capture DISTRIBUTIVE inferences that arise in the context of nominal quantifiers.

$$S^* = \begin{Bmatrix} \{w_A\} & \mapsto A \\ \{w_B\} & \mapsto B \\ \{w_{AB}\} & \mapsto A \wedge B \\ \{w_A, w_{AB}\} & \mapsto A, A \vee B \\ \{w_B, w_{AB}\} & \mapsto B, A \vee B \\ \{w_A, w_B\} & \mapsto A \vee B \\ \{w_A, w_B, w_{AB}\} & \mapsto A \vee B \end{Bmatrix} \quad R^* = \begin{Bmatrix} A & \mapsto \{w_A\} \\ B & \mapsto \{w_B\} \\ A \wedge B & \mapsto \{w_{AB}\} \\ A \vee B & \mapsto \{w_A, w_B\} \end{Bmatrix}$$

Fig. 6 Optimal strategies for sender and receiver in a lifted IBR model in Franke (2009)

A reviewer observed that, from a classical Gricean perspective, implicatures involve pragmatic inferences about the speaker's information state. Consequently, one might argue that there is no reason to detect implicatures in the current experimental design, since the listener (i.e., the participant) is fully aware of the speaker's information state (i.e., the fictional character). Therefore, it should not be surprising that UNCERTAINTY inferences are not derived. Firstly, if this argument holds, we would also expect not to find POSSIBILITY inferences (unless such inferences have another non-Gricean source). However, the mean acceptance rate of our TARGET-2 condition clearly goes against this expectation. Secondly, the majority of verification tasks in the literature assumes that full information is available to both speaker and listener. For instance, in the case of scalar items like *some*, it has been consistently observed that in contexts where the stronger alternative *all* is true, sentences with *some* are still associated with lower judgments of acceptability in a truth-value task, which we would not expect if no scalar inference is generated.

A promising approach to interpreting our findings involves game-theoretic models of pragmatics, which model interactions between speaker and listener roles explicitly. Game-theoretic models have gained traction as tools for capturing the dynamic strategies of communication (e.g., Dekker and van Rooy 2000; Benz et al. 2005; Franke 2009; Frank and Goodman 2012). By formalizing speaker and listener strategies, these models analyze how inferences emerge from the cooperative (or strategic) interplay between conversational partners. We assess the compatibility of one such model, the Iterated Best Response (IBR) model proposed by Franke (2009), with our experimental results. Although other recent models could and should also be considered (e.g., Champollion et al. 2019; Franke and Bergen 2020; Cremers et al. 2023), the IBR model offers a compelling parallel to the notion of *state* in Aloni's 2022 non-implicature approach discussed in Sect. 7.3, and it will allow us to distinguish speaker and listener roles in inference derivation.

The IBR model defines a set of states (sets of worlds) and possible messages that speakers may convey. The speaker (sender) selects an utterance (message) that conveys the intended meaning, while the listener (receiver) interprets it based on some optimal strategies. We assume the possible messages here include A, B, the disjunction $A \vee B$, and the conjunction $A \wedge B$. The speaker and listener are modelled as rational agents who aim to maximize communicative utility within given constraints.

Figure 6 displays the optimal strategies for the sender (S^*) and the receiver (R^*). The sender sends a message based on a certain state. Importantly, the states where a disjunction $A \vee B$ is possible pattern with the results discussed in Sects. 4 and

5, as only states where POSSIBILITY is true are admitted. The receiver interprets the message of the sender, and a disjunction is associated with a state where both UNCERTAINTY and EXCLUSIVITY need to hold.[25]

This discussion leads to two observations. First, as said, speakers are not using a disjunction literally, and only states compatible with POSSIBILITY are possible, similarly to the behavior of the participants in our experiments. Sentence-picture verification tasks, like those in our study, are often seen as closer to production measures, rather than interpretation (Degen and Goodman 2014). Consequently, we conjecture that in a sentence-picture verification paradigm, participants reason from the point of view of the sender/speaker: given what they know (i.e., their information state), they determine whether a certain statement can be asserted. There are of course cases where UNCERTAINTY and EXCLUSIVITY should be drawn. The IBR model discussed above leads to the conjecture that these additional inferences arise on the receiver/listener end: given an utterance, a listener determines the most optimal interpretation for that utterance. This would suggest that while POSSIBILITY is in this sense non-conversational, UNCERTAINTY and EXCLUSIVITY emerge from the dynamics of conversation between speaker and hearer.

While the IBR theoretic model above highlights an interesting perspective on the speaker vs. hearer distinction and our experimental results, more attention should be paid to recent developments in the field of game-theoretic pragmatics. In particular, it is important to take into consideration more recent game-theoretic approaches (Champollion et al. 2019; Franke and Bergen 2020; Cremers et al. 2023; Alsop 2024) and explore the integration of the QUD into the model, as discussed at the end of Sect. 7.2.3, as well as the possible integration of the factors responsible for the derivation of POSSIBILITY in the non-implicature approaches (e.g., neglect-zero, homogeneity, …) into game-theoretic models.

The next section then briefly overviews contexts where UNCERTAINTY and EXCLUSIVITY inferences could be drawn and their interaction with POSSIBILITY.

7.5 EXCLUSIVITY and UNCERTAINTY

Our experiments showed that disjunction could give rise to POSSIBILITY, while UNCERTAINTY and EXCLUSIVITY were not derived. Still, it is natural to assume that in certain contexts those inferences that we typically associate with disjunction are generated. In line with what was observed at the end of the last section, we conjecture that an experiment aimed at measuring interpretation rather than production could provide valuable insights. In the current mystery box experiments, participants had to judge if a sentence was good or bad based on the information available to them. A variant of these experiments would consist in asking participants to correctly place the balls inside the boxes given a truthful sentence, thereby reflecting the listener/interpretation viewpoint.

[25] We note that the IBR model above relies on two fundamental assumptions (Franke 2009, pp. 143-145). First, speakers are competent (i.e., they consider more uncertain states less likely). Second, speakers will select only the most likely messages in a state of maximal information (i.e., not a singleton set), while they will be indifferent when the state is not of maximal information. We leave to further research a more comprehensive discussion of these modeling assumptions.

Another case in which EXCLUSIVITY and UNCERTAINTY seem to play a role are conversational contexts like (32) and (33). In such contexts, a disjunction results in oddity. A preliminary explanation could attribute the oddness of such continuations to the fact that EXCLUSIVITY and UNCERTAINTY inferences clash with the information available in the context. In this regard, an important question is whether UNCERTAINTY inferences always correlate with EXCLUSIVITY. The seeming oddness of (33-b) suggests a negative answer.

(32) Context: We all agree that Mary is here and John is here as well.
  a. ?Mary or John are here.

(33) Context: We all agree that Mary is here, and John might be too.
  a. ?Mary or John are here.
  b. (?)I don't know whether they are both here, but either one or the other is here.

An alternative explanation for the oddness observed in cases such as (32) and (33) is that we are bringing attention to the alternatives in a linguistic form. For example, in (33), we are highlighting the alternative that 'It is certain that Mary is here and it is possible John is here'. This alternative is generally more complex, and it is typically not considered unless we emphasize it linguistically, as suggested by structure-sensitive characterizations of alternatives like Katzir (2007). We leave the investigation of cases like the above for future work.

8 Conclusion and next steps

This study has presented empirical evidence challenging the traditional implicature-based account of IGNORANCE inferences linked to plain disjunctive sentences. Through two experiments, we showed that participants readily derived POSSIBILITY inferences independently of UNCERTAINTY inferences, casting doubt on the classical view of these inferences and supporting alternative frameworks. This work thereby contributes to ongoing debates in semantics and pragmatics concerning the nature of disjunctive inference patterns and suggests that a reevaluation of classical implicature-based accounts is warranted.

Our discussion contextualized these findings within three theoretical perspectives: grammatical, non-implicature, and, briefly, game-theoretic approaches. Notably, our results align closely with non-implicature accounts that propose that POSSIBILITY inferences are independent and need not be contingent on UNCERTAINTY inferences. Recent implicature accounts can also account for our results. However, they introduce additional mechanisms of inference derivation. Non-implicature accounts, by contrast, provide a more parsimonious explanation yet still require a principled mechanism for deriving scalar implicatures.

Several directions for further research could deepen our understanding of these inferences and their status at the semantics-pragmatics interface. Future studies could employ diverse experimental methodologies, such as indirect judgments of felicity, to test the robustness of our findings across tasks. We have highlighted the potential of QUD-based explanations: theoretically, by contributing to models that explicitly integrate the QUD in inference derivation, and experimentally, by designing studies that can rigorously control for QUD effects. Additionally, further exploration of contexts discussed in Sect. 7.5, where UNCERTAINTY and EXCLUSIVITY interact, may yield novel insights into the mechanisms behind these inferences.

Other promising directions include investigating how contextual and cooperative factors influence the availability of these inferences. For instance, in mathematical reasoning contexts, disjunction might be interpreted literally, thereby minimizing the derivation of both POSSIBILITY and UNCERTAINTY inferences. Relatedly, non-cooperative settings, where conversational maxims are relaxed, could shed light on the source of these inferences. Finally, our results reveal individual variation, suggesting that some participants may be inherently more inclined to derive POSSIBILITY inferences from disjunction than others. Investigating the factors driving these differences presents an interesting avenue for future work.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11050-024-09226-3. We have included a detailed description of the experimental instructions as supplementary material to facilitate replicability and ensure transparency of the experimental design.

Acknowledgements We would like to the audiences of SuB 28, XPRAG X, the NihiL workshop, TLLM IV for valuable questions and comments. We are very grateful to the reviewers and the editor, Florian Schwarz. We thank Luka Crnič, Milica Denić, Patrick Elliott, Danny Fox, Daniel W. Harris, Alex Kalomoiros, Roni Katzir, Matt Mandelkern, Andreea Nicolae, Rick Nouwen, Floris Roelofsen, Benjamin Spector and Raven Zhang for discussion.

Authors' Contribution Statement All authors contributed to the study conception. The experimental design was adapted from a previous experimental paradigm used in Marty et al. (2023). Material preparation, data collection were performed by PM. Data analysis was done by PM in coordination with MD, SR, and JR. The current draft of the manuscript was written by MD (Sects. 1, 2, 3, 6, 7, 8) and PM (Sects. 3, 4, 5, 6). PM wrote Sects. 4.6, 5.6 and 6.6 in coordination with MD. MA and JR provided feedback on the first draft. JR collaborated with MD and PM to produce revised drafts for the other authors to review. All authors edited and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Experiment 1, Experiment 2 and Experiment 3 were funded by the ILLC-Formal Semantics and Philosophical Logic Unit, University of Amsterdam.

Data and Code Availability Stimuli, data and analysis code associated with the experiments are available open access on the OSF platform at https://osf.io/4ut2c/.

Declarations

Ethics Approval All experiments involved human participants. All studies have been approved by the Ethics Committee of Faculty of Humanities at the University of Amsterdam. Data were collected, anonymized and stored according to GDPR guidelines.

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

Aloni, M. 2022. Logic and conversation: The case of free choice. *Semantics and Pragmatics* 15:5. https://doi.org/10.3765/sp.15.5.

Aloni, M., and P. van Ormondt. 2023. Modified numerals and split disjunction: The first-order case. *Journal of Logic, Language and Information* 32:539–567.

Alsop, A. 2024. The pragmatics of free choice *any*. *Glossa: a Journal of General Linguistics* 9(1):1–36. https://doi.org/10.16995/glossa.9629.

Anwyl-Irvine, A. L., J. Massonié, A. Flitton, N. Kirkham, and J. K. Evershed. 2020. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods* 52:388–407.

Bar-Lev, M. E., and D. Fox. 2020. Free choice, simplification, and innocent inclusion. *Natural Language Semantics* 28(3):175–223.

Bar-Lev, M. E., and D. Fox. 2023. On fatal competition and the nature of distributive inferences. *Natural Language Semantics* 31(4):315–348. https://doi.org/10.1007/s11050-023-09210-3.

Barr, D. J., R. Levy, C. Scheepers, and H. J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3):255–278. https://doi.org/10.1016/j.jml.2012.11.001.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48. https://doi.org/10.18637/jss.v067.i01.

Benz, A., G. Jäger, and R. van Rooij. 2005. *Game theory and pragmatics*. London: Palgrave Macmillan.

Buccola, B., and A. Haida. 2019. Obligatory irrelevance and the computation of ignorance inferences. *Journal of Semantics* 36(4):583–616. https://doi.org/10.1093/jos/ffz013.

Champollion, L., A. Alsop, and I. Grosu. 2019. Free choice disjunction as a rational speech act. In *Proceedings of SALT 29*, eds. K. Blake, et al., 238–257.

Chierchia, G., D. Fox, and B. Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In *Handbook of semantics: An international handbook of natural language meaning*, eds. P. Portner, C. Maienborn, and K. von Heusinger, 2297–2332. Berlin: Mouton de Gruyter.

Cremers, A., L. Coppock, J. Dotlačil, and F. Roefolsen. 2022. Ignorance implicatures of modified numerals. *Linguistics and Philosophy* 45(3):683–740.

Cremers, A., E. G. Wilcox, and B. Spector. 2023. Exhaustivity and anti-exhaustivity in the RSA framework: Testing the effect of prior beliefs. *Cognitive Science* 47(5):e13286. https://doi.org/10.1111/cogs.13286.

Crnič, L., E. Chemla, and D. Fox. 2015. Scalar implicatures of embedded disjunction. *Natural Language Semantics* 23(4):271–305. https://doi.org/10.1007/s11050-015-9116-x.

Degen, J., and N. Goodman. 2014. Lost your marbles? The puzzle of dependent measures in experimental pragmatics. In *Proceedings of the annual meeting of the Cognitive Science Society*, 397–402.

Dekker, P., and R. van Rooy. 2000. Bi-directional optimality theory: An application of game theory. *Journal of Semantics* 17(3):217–242. https://doi.org/10.1093/jos/17.3.217.

Dorst, K., and M. Mandelkern. 2022. Good guesses. *Philosophy and Phenomenological Research* 105(3):581–618.

Fox, D. 2007. Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, eds. U. Sauerland and P. Stateva, 71–120. London: Palgrave. https://doi.org/10.1057/9780230210752_4.

Fox, J., and S. Weisberg. 2019. *An R companion to applied regression*, 3rd edn. Thousand Oaks: Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Frank, M. C., and N. D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science* 336(6084):998–998.

Franke, M. 2009. Signal to act: Game theory in pragmatics. PhD dissertation, University of Amsterdam. https://eprints.illc.uva.nl/id/eprint/2081.

Franke, M., and L. Bergen. 2020. Theory-driven statistical modeling for semantics and pragmatics: A case study on grammatically generated implicature readings. *Language* 96(2):e77–e96. https://doi.org/10.1353/lan.2020.0034.

Gamut, L. T. F. 1991. *Logic, language and meaning*. Chicago: University of Chicago Press.

Gazdar, G. 1979. *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.

Geurts, B. 2010. *Quantity implicatures*. Cambridge: Cambridge University Press.

Goldstein, S. 2019. Free choice and homogeneity. *Semantics and Pragmatics* 12:23. https://doi.org/10.3765/sp.12.23.

Grice, P. 1975. Logic and conversation. In *The logic of grammar*, eds. D. Davidson and G. Harman, 64–75. Encino: Dickenson.

Grice, P. 1989. *Studies in the way of words*. Cambridge: Harvard University Press.

Harrell, F. E. Jr. 2021. Package 'Hmisc'. R package version 4.5-0. https://CRAN.R-project.org/package=Hmisc.

Hintikka, J. 1962. *Knowledge and belief: An introduction to the logic of the two notions*. Ithaca: Cornell University Press.

Horn, L. 1972. On the semantic properties of logical operators in English. PhD dissertation, UCLA.

Incurvati, L., and J. J. Schlöder. 2019. Weak assertion. *Philosophical Quarterly* 69(277):741–770.

Katzir, R. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30(6):669–690. https://doi.org/10.1007/s10988-008-9029-y.

Levinson, S. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lüdecke, D. 2023. sjPlot: Data visualization for statistics in social science. R package version 2.8.14. https://CRAN.R-project.org/package=sjPlot.

Mandelkern, M., and K. Dorst. 2022. Assertion is weak. *Philosophers' Imprint* 22:19. https://doi.org/10.3998/phimp.1076.

Marty, P., S. Ramotowska, R. Breheny, J. Romoli, and Y. Sudo. 2023. On the source of distributive inferences. Unpublished manuscript. https://lingbuzz.net/lingbuzz/007623.

Marty, P., J. Romoli, Y. Sudo, and R. Breheny. 2024. What makes linguistic inferences robust? *Journal of Semantics* 41(1):1–52. https://doi.org/10.1093/jos/ffad010.

Meyer, M. C. 2013. Ignorance and grammar. PhD dissertation, MIT. http://dspace.mit.edu/handle/1721.1/7582.

Nieder, A. 2016. Representing something out of nothing: The dawning of zero. *Trends in Cognitive Sciences* 20(11):830–842.

Noveck, I. A. 2001. When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition* 78(2):165–188. https://doi.org/10.1016/S0010-0277(00)00114-1.

Palan, S., and C. Schitter. 2018. Prolific.ac — a subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17:22–27.

R Core Team. 2021. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sauerland, U. 2004a. On embedded implicatures. *Journal of Cognitive Science* 5(1):107–137.

Sauerland, U. 2004b. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3):367–391. https://doi.org/10.1023/B:LING.0000023378.71748.db.

Sauerland, U. 2005. The epistemic step. *Experimental Pragmatics* 10:225–250.

Singh, R. 2008. On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy* 31:245–260. https://doi.org/10.1007/s10988-008-9038-x.

Authors and Affiliations

Marco Degano[1] · Paul Marty[2] · Sonia Ramotowska[1] · Maria Aloni[1] · Richard Breheny[3] · Jacopo Romoli[4] · Yasutada Sudo[3]

✉ M. Degano
m.degano@uva.nl

P. Marty
pmarty@edu.ulisboa.pt

S. Ramotowska
s.ramotowska@uva.nl

M. Aloni
m.d.aloni@uva.nl

R. Breheny
r.breheny@ucl.ac.uk

J. Romoli
jacopo.romoli@hhu.de

Y. Sudo
y.sudo@ucl.ac.uk

[1] University of Amsterdam, Amsterdam, The Netherlands

[2] Universidade de Lisboa, Lisbon, Portugal

[3] University College London, London, UK

[4] Heinrich-Heine Universität Düsseldorf, Düsseldorf, Germany