Cognition 244 (2024) 105667

Contents lists available at ScienceDirect

Cognition

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

Full length article

Implicature priming, salience, and context adaptation☆

Paul Marty[a], Jacopo Romoli[b], Yasutada Sudo[c,*], Richard Breheny[c]

[a] L-Università ta' Malta, Malta
[b] Heinrich-Heine-Universität Düsseldorf, Germany
[c] University College London, United Kingdom

ARTICLE INFO

Dataset link: https://osf.io/6gsv9/

Keywords:
Quantity implicatures
Alternatives
Salience
Relevance
Priming effects
Inverse preference
Context adaptation

ABSTRACT

Recent experimental research has observed two kinds of priming effects on quantity implicatures. One is the Strong–Weak contrast, where more quantity implicatures are observed after prime trials forcing interpretations with quantity implicatures ('Strong primes') than after prime trials forcing interpretations without quantity implicatures ('Weak primes'). The other effect is the Alternative-Weak contrast, where prime trials mentioning alternative expressions ('Alternative primes') similarly lead to more quantity implicatures. It has been claimed that both of these effects should be understood in terms of increased salience of alternative expressions used to compute quantity implicatures. We present experimental evidence that speaks against this hypothesis. With the help of novel baseline conditions, which were absent in previous studies on implicature priming, we observe that the results in the priming paradigm commonly used in the literature are *inverse preference effects* in the sense that robust priming effects are observed towards interpretations that are normally unexpected, and depending on the baseline expectation, each of the three prime types mentioned above may have priming effects. We furthermore investigated different types of alternative priming for so-called *ad hoc* implicatures and found that for these implicatures, presenting an alternative expression in a simple sentence does not have a priming effect on the implicature of a similarly simple sentence, but presenting it in a more complex conjunctive construction does. Our results also show that conjunctions of similar but irrelevant expressions have a similarly robust priming effect and that conjunctive sentences with two conjuncts do not give rise to priming effects on the interpretation of sentences of the same syntactic complexity, but those with three conjuncts do. To make sense of these observations, we propose that what crucially matters for priming implicatures is incremental change in one's probabilistic expectations about the current conversational context brought about by a process we call *context adaptation*.

1. Introduction

Experimental pragmatics applies the insights of formal and philosophical analysis of language use in the development of models of the cognitive processes underpinning this human activity. Notably, early research in Noveck (2001) shed light on aspects of child language development with an experimental paradigm based on ideas first shared in Oxford common rooms in the mid-Twentieth Century (Grice, 1975; Wilson & Sperber, 1986). A richer picture of human pragmatic abilities has since been experimentally developed to include links to social-cognitive abilities (Catani & Bambini, 2014; Southgate, Chevallier, & Csibra, 2010; Spotorno, Koun, Prado, van der Henst, & Noveck, 2012), executive function (Antoniou, Cummins, & Katsos, 2016), as well as linguistic-semantic knowledge (Huang & Snedeker, 2009a; Noveck, Chierchia, Chevaux, Guelminger, & Sylvestre, 2002). Increasingly, experimentally-supported models of pragmatics are cast in probabilistic terms (Bergen, Levy, & Goodman, 2016; Catani & Bambini, 2014; Frank & Goodman, 2012; Franke & Jäger, 2016).

While linguists's and philosophers' analyses of pragmatic phenomena are often the inspiration for experimental pragmatics research, experimental work can also help decide among different analyses of linguistic phenomena, such as the nature of presupposition projection (Chemla, 2009; Schwarz, 2007; Schwarz & Tiemann, 2017), numerical quantification (Marty, Chemla, & Spector, 2015), plurality (Maldonado, Chemla, & Spector, 2017, 2019), etc. In the limiting case, the same theoretical construct appears both in linguistic analysis and psychological models of pragmatic abilities and that construct can become the subject of experimental research. In this paper, we focus on one such notion, *salience*, which plays an important role both in the

☆ This paper is a part of special issue "20 Years of XPrag".

* Corresponding author.
E-mail address: y.sudo@ucl.ac.uk (Y. Sudo).

https://doi.org/10.1016/j.cognition.2023.105667
Received 31 October 2022; Received in revised form 4 October 2023; Accepted 15 November 2023
Available online 4 January 2024

theory of quantity implicatures and in psycho-linguistic models of how quantity implicatures are processed.

In the following section, we outline the standard picture of how quantity implicatures are processed. We then review theoretical, psycholinguistic and developmental literature which is the site of a current debate about the role for salience in this process. In what follows, we present three priming experiments which provide clear tests of salience-based hypotheses. Our claim is that such hypotheses fall short of a full account of our results and that a context-adaptation approach to quantity implicatures offers a better fit for the picture which emerges.

2. The relevance and salience of alternatives for quantity implicatures

Expressions like *some* are often associated with inferences called *quantity implicatures*. One characteristic of implicatures is that they are not always present, which suggests that they are not directly encoded in the core semantics of the expressions used. To illustrate, consider the following sentence:

(1) Some of the symbols are circles.

a. All of the symbols are circles. ALTERNATIVE
b. Not all of the symbols are circles. IMPLICATURE

One often understands such a sentence as implying that not all of the symbols are circles, but this inference is considered to be a quantity implicature (or more specifically a *scalar implicature*), as it is not always present. This is evidenced by the consistency of *Some of the symbols are circles, and in fact, all of them are*, for example. Furthermore, embedding *some* in certain grammatical contexts makes it natural to not have the inference. Concretely, consider the question *Are some of the symbols circles?*. A negative answer to this question would most naturally be understood as meaning that none of the symbols are circles, rather than that none or all of the symbols are circles, which is what it should mean if it were the negation of the some-but-not-all meaning.

Facts like above have been argued to show that the core meaning of *some* does not include the quantity implicature 'not all', but if so, why do we often draw this inference, and when do we do so? A number of different answers to these questions have been offered in the theoretical literature, but one common set of assumptions is that their derivation crucially involves (i) referring to an alternative expression to what is uttered—or simply an *alternative*—and (ii) negating that alternative. More specifically, it is most commonly assumed that an alternative is a linguistic expression that is distinct from, but related in some way to the expression used.[1] For instance, for *Some of the symbols are circles*, the crucial alternative is taken to be *All of the symbols are circles*, whose negation amounts to the quantity implicature that not all of the symbols are circles.

Theories of quantity implicatures often assume that alternatives must be contextually relevant in some way for quantity implicatures to arise, e.g., by virtue of addressing a question raised in the current utterance context (a.o., Fox & Katzir, 2011; Geurts, 2010). To illustrate this point concretely, suppose that you ask someone *What symbols are on that card over there?*, and they answer, *Some of the symbols are circles*. Intuitively, the quantity implicature that not all of the symbols are circles is perceived very robustly. Compare this to a different scenario where you ask someone *Are there any circles on that card over there?*, and they answer by uttering *Some of the symbols are circles*. In this case, the quantity implicature feels much less robust, and the reply could be understood as being open about whether or not all of them are circles. When an utterance of an expression is able to address a question, whether explicitly raised or implicit in the context, the information encoded in the expression, as well as the expression itself, is described as being relevant to the 'Question under Discussion' (QuD, cf. Roberts, 2012). We will make heavy use of this notion of contextual relevance below, for ease of exposition. But it should be noted that a better picture of how information is or is not relevant will likely take us beyond a simple QuD-based model (see Büring, 2003; Cremers, Wilcox, & Spector, 2023; Van Rooy, 2003).

An important and yet open question about the theory of quantity implicatures concerns which expression should count as an alternative to a given expression. The following consideration shows that this is not a trivial question. Positing (1a) as the relevant alternative to (1) explains the implication in (1b), but what about an equally relevant alternative to (1), like *Some and not all of the symbols are circles*? The negation of this hypothetical alternative would lead to the unattested implication that all of the symbols are circles, so we would like to rule it out as a possible alternative (Breheny, Klinedinst, Romoli, & Sudo, 2018; Fox, 2007; Katzir, 2007). To resolve this issue—often called the *symmetry problem*—one might consider a constraint that bans alternatives that are structurally more complex than the asserted sentence (Atlas & Levinson, 1981), but cases where quantity implicatures seem to involve structurally more complex alternatives have been raised (Matsumoto, 1995). For this reason, Katzir (2007) put forward a more sophisticated view, according to which alternatives can be no more complex in their linguistic structure *unless the linguistic structure has been made salient in the discourse* in some way, e.g., by virtue of having been recently used (see also Fox & Katzir, 2011).[2]

Thus, it is widely considered in the current literature on quantity implicatures that both salience and contextual relevance may play a role in determining which alternatives to negate in the computation of quantity implicature. Importantly, we should point out that salience and contextual relevance are considered as independent notions: a relevant alternative may not be immediately salient in discourse and a salient alternative need not be relevant to the task at hand. An example of the second case is given in (2) (adapted from Romoli, 2012).

(2) I don't know whether all of the symbols are circles. But some of them are.

In this example, the second sentence with *some* does not give rise to a 'not all' implicature, despite the alternative *All of the symbols are circles* being salient in the previous discourse. This is arguably because such an alternative is not relevant here, given that the speaker has just signalled that they are not in a position to answer whether all of the symbols are circles.

3. Variation of implicature rates, salience and relevance

Previous quantitative studies on implicatures have established that how often quantity implicatures are observed is dependent on the experimental task, and that different expressions give rise to quantity implicatures to different degrees of robustness even with respect to the same experimental method (e.g., Geurts & Pouscoulous, 2009; Pankratz & van Tiel, 2021; van Tiel, van Miltenburg, Zevakhina, & Geurts, 2016; van Tiel, Pankratz, Marty & Sun, 2019; van Tiel, Pankratz & Sun, 2019).

For example, Bott and Noveck (2004) report that around 60% of participants respond 'False' to critical items like *Some elephants are mammals*, meaning that those participants understand the sentence to mean some and not all elephants are mammals, while 40% of those

[1] Some scholars take alternatives to be alternative 'meanings' rather than alternative expressions (e.g., van Rooij & Schulz, 2004, 2006) and there are even theories that attribute at least some cases of quantity implicatures to the conventional meaning of expressions uttered, (e.g., Fine, 2017). Our final conclusion in this paper will not hinge on the assumption that alternatives are linguistic expressions, but we tentatively assume so for the ease of exposition.

[2] An alternative theory to Katzir's has been developed within Bayesian models of quantity implicature (Bergen et al., 2016; Cremers et al., 2023). Here we do no attempt to delve into this debate itself, as our focus is exclusively on the notion of salience.

tested understand the sentence without the implicature. On the other hand, Papafragou and Musolino (2003) asked a control group of adults to determine whether sentences like *Some of the horses jumped over the fence* are true or false against a scene played out before them, and found that the rejection rate was much higher, at around 90%. Geurts and Pouscoulous (2009) compared two tasks with respect to how people understand sentences like *Some of the B's are in the box on the left*. One task was a verification task against pictures, and the other task was an inference task of judging if the sentence implies 'Not all of the B's are in the box on the left'. They report far higher implicature-based responses in the inference task than the verification task (62% vs. 34%). A plausible explanation for why the tasks used in these studies yielded different response rates for quantity implicatures is because they differed with respect to salience and/or relevance of the crucial alternatives. Specifically, Papafragou and Musolino presented a scenario where the speaker's aim and the relevance of their utterance are clear. This contrasts with the studies reported in Bott and Noveck (2004), where stimuli are presented without further context and so there ought to be uncertainty in participants' minds about any imagined relevance. Similarly, the visual salience of the horses in Papafragou and Musolino's study could have led to the increased salience of an alternative way of describing the scenario that involves *all*. When it comes to Geurts and Pouscoulous' comparison, the authors themselves observe that the stimuli in their inference task not only mention the alternative expression, but also may suggest that it is relevant (see also Sun & Breheny, 2022), and argue that these are reasons for the higher rate of implicatures in the inference task compared to the verification task.

Questions about the roles of salience and contextual relevance have also arisen in research on children's ability to derive quantity implicatures. Papafragou and Musolino's study, mentioned above, demonstrates a widely replicated result that adults derive quantity implicatures for *some* at a much higher rate than children. It seems reasonable to assume that one important factor that contributes to this disparity is that children have more limited processing capacities, when it comes to a complex pragmatic inference like implicatures (Huang & Snedeker, 2009b; Pouscoulous, Noveck, Politzer, & Bastide, 2007). To the extent that this is correct, the question arises whether cognitive limitations impact on the salience of alternatives, or establishing relevance. Barner, Brooks, and Bale (2011) come down on the side of alternative salience, on the basis of experimental evidence showing that so-called *ad hoc* implicatures (see below for examples and discussion) are derived even when the lexical 'not-all' implicature for *some* is not. The proposal is that, in their study, the *ad hoc* alternatives (concerning three animal characters present in the display) are highly salient, whilst the lexical alternative, *all* for *some* may not be as strongly associated in the children's memory as it is for adults. As a counterpoint to this proposal, Skordos and Papafragou (2016) propose that the efficacy of salient linguistic information may result from how it affects the certainty about the source of contextual relevance; and this may be responsible for children's improved performance with *ad hoc* implicatures compared to controls. They provide evidence that the salience of a quantifier that cannot directly serve as an alternative to *some*, namely, *none*, also has the effect of improving performance. Skordos and Papafragou reason that sentences with *none*, like sentences with *all* can both have the effect of promoting the right kind of QuD in children's minds to make the quantity implicature easier to derive.

To sum up this discussion, the idea that linguistic alternatives must be salient for quantity implicatures to arise plays an important role in theoretical accounts. Equally, it is generally agreed that quantity implicatures do not arise if the alternative is not relevant in the context. We highlighted that it is possible that either salience of alternative or contextual relevance may be factors in explaining differences in an individual's propensity to incorporate a quantity implicature in a given situation. As we explain below, we see a recently developed priming paradigm for quantity implicatures as providing a helpful environment in which to make a more controlled investigation of the roles of salience and contextual relevance in implicature derivation.

4. Priming research and salience in implicature variability

Several recent studies have demonstrated that rates of implicature-based responses are systematically affected by manipulations in a priming phase. According to one interpretation of these outcomes, salience of alternative is an important factor in increasing rates of implicature response. A key initial set of results was reported in Bott and Chemla (2016), which used a priming paradigm based on Raffray and Pickering (2010). Since our experiments below, like other recent studies on implicature priming, adopt Bott and Chemla's design, it is worth reviewing their paradigm in some detail at this point. Bott and Chemla employed a picture selection task for their critical trials, where each trial had two pictures, one overt and one hidden, the latter with a label 'Better Picture?', as illustrated in Fig. 1(a). These pictures were presented together with a sentence that can potentially have a quantity implicature, such as, *Some of the symbols are squares*. Crucially, the overt picture is only compatible with the reading of this sentence without the implicature, so that the participant would choose the overt picture only if they considered this weaker reading to be acceptable, and choose the covered card if they feel that the sentence carries the implicature. On this linking assumption, a covered card choice stands as a proxy measure for accessing the implicature-based reading.

Each covered-card trial of Bott and Chemla's experiments was preceded by two priming trials. Unlike target trials, both pictures in prime trials are visible, as illustrated in Figs. 1(b) and 1(c). There are two types of priming trials, weak primes and strong primes. In a weak prime, only one of the two pictures makes the sentence true while the other one renders the sentence clearly false. Crucially, the picture that makes the sentence true does so only if it is understood without an implicature. We shall call this a 'weak reading' for the sentence. For example, in weak prime trials, the true card for the sentence *Some of the symbols are crosses* shows that all the symbols are crosses, while the other card clearly falsifies the sentence, i.e., none of the symbols are crosses (see Fig. 1(b) for example). Consequently, the participant is forced to choose the first picture, which in turn means that they have to access the weak, without-implicature reading. In a strong prime, the sentence can be true with respect to both of the overt pictures. One of the pictures is the same as in the corresponding weak prime, and the sentence is only true if it is understood on its weak reading. Crucially, the second picture is one where the sentence is true on both weak and strong readings (see Fig. 1(c) for example). Participants are asked to choose the best match for the sentence and typically choose the picture that makes both readings of the sentence true.

Bott and Chemla used three types of linguistic stimuli: sentences with *some* (as in Fig. 1), sentences with numerals and sentences that could trigger *ad hoc* implicatures, more details of which will be given below.[3] They found that, for all three expression types, the hidden picture was chosen more often in target trials following strong primes than in target trials following weak primes. This difference between weak and strong prime conditions was consistently replicated by later studies (Meyer & Feiman, 2021; Rees & Bott, 2018; Waldon & Degen, 2020). Furthermore, a difference in the same direction was also observed when the target trial involved a different expression type than the priming trials ('cross-scale' priming), albeit the size of the difference was considerably smaller than when the target and priming trials involved the same expression type (see also Meyer & Feiman, 2021).

The fact that cross-scale priming was observed led Bott and Chemla to argue that this provides evidence that there is a common mechanism behind the relevant inferences of the three expression types, which can be primed in the experimental paradigm under discussion. Recall

[3] In fact, in their Experiment 3, Bott and Chemla (2016) also tested a fourth type of expression, bare plurals. Our experiments, as well as other recent priming studies do not include this expression type, so we will ignore it here.

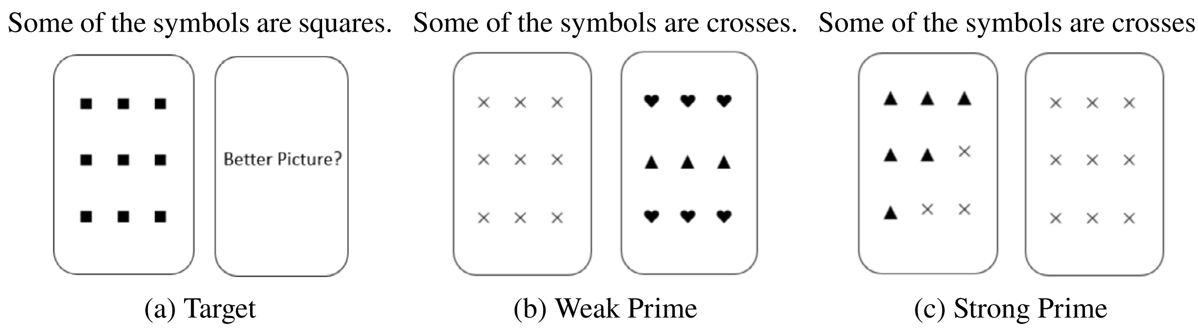

Fig. 1. Example items illustrating the logic behind the SOME trials in Bott and Chemla (2016) (these are not the actual items they used).

that, generally, the derivation of a given implicature involves two steps: (i) referencing an alternative and (ii) negating it. Since different scales involve different alternatives, what is common across scales must be the mechanism that negates alternatives, i.e., step (ii) above. Bott and Chemla further hypothesise that the considerably larger priming effect observed for within-scale priming is due to priming the use of a particular alternative. Specifically, they claim that a strong prime forces the participant to find the alternative and negate it, i.e., steps (i) and (ii) above. If we assume that the negation mechanism gives rise to a small priming effect as in the case of cross-scale priming, it must be that referencing an alternative significantly boosts the salience level of that alternative, thereby making it more likely to be used when the same scalar item is encountered afterwards (for related claims, see Rees & Bott, 2018; Waldon & Degen, 2020). In this way, Bott and Chemla account for the larger effect size of within-scale priming as a combined effect of two types of priming. We refer to an account of the large boost in within-scale priming through activation of alternatives as the Salience Hypothesis.

Rees and Bott (2018) argue for the Salience Hypothesis based on data from priming experiments which include primes formed using the alternative sentences themselves. The paradigm closely followed that in Bott and Chemla (2016) where each of the target trials was preceded by two priming trials of the same kind. Rees and Bott (2018) tested three kinds of priming: Strong, Weak and Alternative, as illustrated in Fig. 2. Strong and Weak primes were the same as in Bott and Chemla (2016). Alternative primes used sentences containing the crucial alternative expression that is to be used to draw the implicature in the target trial following them. So, in cases where the target sentence involved the quantifier *some*, the prime trials involved sentences with *all* (see Fig. 2(c) for example). Since the whole purpose of an alternative prime is to present the alternative expression overtly, it does not matter whether or not its interpretation involves an implicature, and consequently the logic behind the pictures is simple: both pictures are overt and one of the pictures simply makes the alternative sentence true, while the other one makes the sentence false.[4]

Rees and Bott (2018) report results indicating that for all three expressions, quantity implicatures were observed more often after Strong primes than after Weak primes, which replicates the finding in Bott and Chemla (2016) and elsewhere. The novel finding in Rees and Bott was that more quantity implicatures were observed after Alternative primes than after Weak primes such that the difference between Strong and Weak primes was comparable to the difference between Alternative and Weak primes. Based on these results, Rees and Bott argue for the Salience Hypothesis as follows. They start with the assumption that the difference between Strong and Weak primes is to be explained in terms of the Salience Hypothesis, as previously proposed by Bott and Chemla (2016). That is, Strong primes make alternatives salient by virtue of forcing implicatures to be computed, and this leads to boosting effects in the following target trials. Now, since Alternative primes had very similar boosting effects, the Salience Hypothesis is enough to explain their effects as well.

However, recent findings cast doubt on Rees and Bott's reasoning. Waldon and Degen (2020) conducted a similar study and observed that the priming effect of the Alternative primes is weaker than that of the Strong primes.[5] Furthermore, they pointed out that the experiments reported in Rees and Bott (2018) do not have baseline conditions and therefore it is not clear whether it is the Strong and Alternative primes that drive the priming effects by boosting implicatures, as Rees and Bott assume, or the Weak primes that actually have inhibition effects, or it could even be that both of these effects are present simultaneously. To this end, Waldon and Degen (2020) included baseline trials, which were target trials preceded by arithmetic problems that were assumed to have no effect on implicature computation. Their results provide weak evidence that Strong primes have boosting effects relative to the baseline while suggesting that Alternative primes actually have no boosting effects. This is contrary to the Salience Hypothesis.

Here we would like to point out a further complication. Although Waldon and Degen's baseline conditions give us some sense of the direction of priming, these baselines were potentially affected by priming trials elsewhere in the experimental session, and not just the immediately preceding neutral prime trials. Previous syntactic priming research has found similar 'spillover effects'. Fine, Jaeger, Farmer, and Ting (2013) account for such effects in terms of how exposure to an unexpected parse results in a larger adjustment to prior expectations than exposure to a commonly encountered parse. In the case of implicature priming, we have reason for concern that something similar may be going on. Suppose that, at some level of description, implicature priming effects result from expectancy adaptation (see Schuster & Degen, 2020; Waldon & Degen, 2020). Suppose also, for the sake of argument, that reading a sentence containing *some* without implicature is not as common as reading it with an implicature. According to what has been learnt in syntactic priming research, exposing participants to Weak and Strong prime trials would result in modulation of expectations in favour of more weak readings. Thus, in the course of experimental sessions containing all priming types, outcomes after neutral primes will reflect a lower expectation for strong readings than a baseline expectation that may be the default in everyday language use, where proportionally fewer weak readings are encountered.

In order to isolate the baselines from priming trials, we will adopt a block design where a Baseline condition will be administered in a first block of covered card trials that involve no priming trials. In a second block of trials, participants will encounter covered card

[4] Rees and Bott (2018) report two versions of the experiment that differed with respect to the incorrect picture for the Weak primes, and the examples in Fig. 2 are closer to their Experiment 2. They observed no essential difference in the results of the two experiments.

[5] Waldon and Degen (2020) also tested priming trials involving different types of 'Exhaustive' alternatives, which expressed the content of the implicature explicitly. They observed certain differences between Alternative and Exhaustive primes, but these differences do not concern us here.

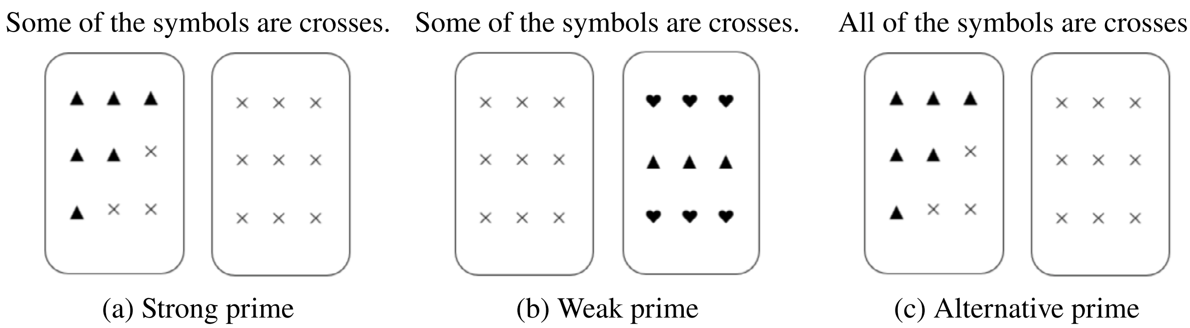

Fig. 2. Example prime trials for SOME in the experiments reported in this paper. Rees and Bott (2018) used items that are identically structured.

trials which follow prime trials. In this way, we can answer more confidently two questions raised about previous implicature priming results: What is the direction of priming? And whether and to what extent do we encounter spillover effects? Before reporting on our first experiment, which explores these two questions, we would like to frame an alternative hypothesis for previous priming results, which can serve as a counterpoint to the Salience Hypothesis, which we call the *Context Adaptation Hypothesis*.

As already mentioned, several previous accounts of priming are cast in terms of how the prime stimulus may impact the normal expectations which we bring to bear when processing the target stimulus (Fine et al., 2013; Jaeger & Snider, 2013; Myslín & Levy, 2016). Our hypothesis makes no reference to alternative linguistic or semantic representations *per se*, but takes its lead from more dynamic approaches, according to which language use has the effect of updating an information state or 'context'. Contexts, in this model, contain information about what is relevant or useful information, as well as background information, and so forth (for more discussion, see Cremers et al., 2023). From this perspective, an alternative to an expression may or may not be relevant to the understood context, depending on what the current QuD is. We make a widely adopted assumption that comprehension processes implement a joint inference about what is the current context and what is the intended interpretation of the sentence uttered (Franke & Jäger, 2016; Kao, Wu, Bergen, & Goodman, 2014). At the beginning of an experiment, or more generally of a linguistic communication, one has some implicit prior expectations, given a linguistic stimulus and other information available, about how likely it is that the targeted context is one in which the asserted proposition leaves open questions in a way that an alternative update would not, and how likely it is that the targeted context is one in which the asserted utterance would yield a quantity implicature. Language users are willing to alter and adjust such expectations to the local environment, making use of any available cues that may come up during the experiment or conversation. This adaptation process is assumed to be rapid and incremental, allowing one to flexibly adjust one's linguistic expectations in an ever changing conversational situation, which helps one carry out the conversation more efficiently. According to this view, prime stimuli can affect the computation of a quantity implicature in subsequent stimuli, on the assumption that they can shift expectations about what is the likely target context.

5. Experiment 1: Implicature priming revisited

Experiment 1 was designed to replicate the results for the Strong, Weak and Alternative conditions tested by Bott and Chemla (2016) and Rees and Bott (2018) while including novel Baseline conditions to compare these priming conditions to. In the current study, we do not test for cross-scale priming but follow on from the design in Rees and Bott (2018) with the crucial difference being the inclusion of Baseline trials. Our first aim was to determine the direction of priming. To this end, we adopted a block design and placed all the baseline trials before all the priming trials so as to avoid potential spillover effects from the latter onto the former. For the picture stimuli, we followed Bott and Chemla (2016) in using shape symbols.

5.1. Data availability

Stimuli, data, and analysis code for Experiment 1 are all freely available on the OSF platform at https://osf.io/6gsv9/.

5.2. Methods

5.2.1. Participants

179 native speakers of English participated in this experiment (108 female, average age 35.2 years). Participants were recruited online through Prolific.ac (https://www.prolific.co; see Palan & Schitter, 2018 for an overview) using the following pre-screening criteria: English as a first language, UK/US IP addresses, minimum 90% prior approval rating. Participants were paid £1.40, and average completion time was about 9 min. Participants gave written informed consent. Data were collected and stored in accordance with the provisions of Data Protection Act 2018, the UK's implementation of the General Data Protection Regulation. The experiment was approved by the Research Ethics Committee at UCL.

5.2.2. Experimental sentences and their alternatives

Like much previous implicature priming research, we considered priming effects on three sentence types which are liable to give rise to quantity implicatures at variable rates. We have already considered in detail the case of sentences with *some*. Following Bott and Chemla (2016), Rees and Bott (2018) and others, we include two other sentence types, NUMBER and AD HOC, exemplified in Table 1.[6]

We aimed to construct Strong and Weak prime stimuli, similar to those found in Bott and Chemla (2016) and elsewhere. In addition, we aimed to construct Alternative prime stimuli in the spirit of Rees and Bott (2018). To clarify what this means, consider the case of *some*. As set out in the introduction, we can assume that to account for the implicature for *some*, the alternative would involve *all*. Similarly, according to standard assumptions, an alternative for sentences with numerals would be the same sentence where the numeral is replaced by a higher numeral. Ad hoc implicatures, on the other hand, involve contextually determined, 'ad hoc' alternatives. We follow previous priming research here in assuming that in the context of our study, alternatives

[6] The strong 'exact' reading associated with numeral expressions like *four* is often considered to result from a quantity implicature derived in reference to other numeral expressions like *five*, *six*, etc. for reasons similar to the case of 'some' discussed above. Thus, we tentatively assume here that it is a kind of quantity implicature. However, certain differences from canonical quantity implicatures are also observed, which has led some scholars to claim that this is actually not a case of quantity implicature (Breheny, 2008; Geurts, 2006; Marty, Chemla, & Spector, 2013). While this debate is relevant in the context of this paper and cannot be ignored, we note that it does not affect the predictions of our Context Adaption Hypothesis, which only requires that numerals be ambiguous in some way, irrespective of how their strong reading actually comes about.

Table 1
Illustrations of the sentence types tested in Experiment 1 with their relevant alternative(s) giving rise, upon negation, to the quantity implicature of interest (see Fig. 3 for example target trials involving these sentences). For AD HOC, the alternative in (i) is logically stronger than the base sentence while the alternative in (ii) is logically independent from it.

SOME	Sentence		Some of the symbols are squares
	Alternative		All of the symbols are squares.
	Implicature		Not all of the symbols are squares.
NUMBER	Sentence		There are four squares.
	Alternative		There are five/six/etc. squares.
	Implicature		There are no more than four squares.
AD HOC	Sentence		There is a square.
	Alternative	(i)	There is a square and a star.
		(ii)	There is a star.
	Implicature		There is no star.

for a sentence like *There is a square* could be formed by replacing the mentioned symbol with others encountered in the experimental session, for example *star*, *heart*, etc. and especially with other symbol types depicted on the symbol cards accompanying the sentence of interest. In the two previous priming studies that tested Alternative primes – namely, Rees and Bott (2018) and Waldon and Degen (2020) – the sentences for Alternative primes involved a conjunction, e.g., *There is a square and a star*, while the target trial involved a non-conjunctive sentence, e.g., *There is a square*. We believe this design choice was based on the theoretical assumption that quantity implicatures can only be drawn from alternatives that are more informative. For the example at hand, the conjunctive sentence of the prime entails the sentence of the target trial, which means that the former is more informative.

However, recent theories assume that quantity implicatures can be drawn from alternatives that are logically independent and thus not necessarily properly more informative (Bar-Lev & Fox, 2020; Fox, 2007; see Breheny et al., 2018 for more discussion and references). For such theories, the implicature of *There is a square* that there is no star, for example, may arise by negating a simple alternative like *There is a star*, rather than a conjunctive alternative like *There is a square and a star*. The theoretical literature contains independent empirical arguments for allowing for quantity implicatures based on logically independent alternatives, but we will put these aside here, as most of them are not entirely convincing (see Breheny et al., 2018 for discussion). However, one crucial conceptual advantage of accommodating this view should be mentioned. It has been acknowledged that a complete theory of quantity implicature must include a theory of alternatives. While there is arguably no complete theory of alternatives yet, it has been standard to assume that there are structural constraints on the space of alternatives. In particular, as already mentioned, theorists have argued that structurally more complex alternatives are generally ignored, unless they are contextually salient (see also Fox & Katzir, 2011; Katzir, 2007; Trinh & Haida, 2015). If this is true, then in a context where no particular expressions have been made salient, ad hoc implicatures from, say, *There is a square*, should not be drawn based on structurally more complex alternatives like *There is a square and a star*, but only on simple alternatives like *There is a star*. Since ad hoc implicatures are intuitively available in such contexts, logically independent alternatives need to be tested for, to the extent that the structural constraint is upheld.

Note that, in the context of priming experiments, Alternative primes, by assumption, make an alternative salient by virtue of presenting it overtly. Therefore, the priming effect of conjunctive alternatives like *There is a square and a star* will be compatible with the structural constraint. However, it is worth checking if simple alternatives like *There is a star* also have priming effects. If only more informative alternatives can be used to generate quantity implicatures, such simple alternatives should not show priming effects. On the other hand, if logically independent alternatives can also be used to generate quantity implicatures, we would expect similar priming effects from them. The present experiment was designed to address this issue by testing two kinds of Alternative prime for AD HOC sentences like *There is a square*, namely conjunctive alternatives like *There is a square and a star* (X&Y-ALT, hereafter) and simpler alternatives like *There is a star* (X-ALT, hereafter), as illustrated in Table 1.[7]

5.2.3. Stimuli design

The task used for this experiment was the covered picture task described in previous sections, which has been used by previous studies on implicature priming (Bott & Chemla, 2016; Meyer & Feiman, 2021; Rees & Bott, 2018; Waldon & Degen, 2020; but see Bott & Frisson, 2022 for a priming study with reaction times as the dependent measure). All items involved a sentence presented above two pictures. Target sentences were constructed according to the three frames in (3) adopted from Bott and Chemla (2016), where the [symbol] term was a noun denoting a symbol type from the following list: arrow, cross, circle, diamond, heart, square, star and triangle.

(3) a. Some of the symbols are [symbol]. SOME
b. There are four [symbol]. NUMBER
c. There is a [symbol]. AD HOC

Pictures consisted of a card containing either symbols, henceforth overt cards, or the text ‘Better Picture?’, henceforth covered cards. Target and control trials consisted of one covered card and one overt card: a strong card in the true control trials, a false card in the false control trials and a weak card in the target trials. Example target trials are given in Fig. 3. Prime trials consisted of two overt cards: WEAK and ALTERNATIVE prime trials involved a weak card and a false card while STRONG prime trials involved a weak card and a strong card. Example prime trials are given in Fig. 4 (STRONG and WEAK) and in Fig. 5 (ALTERNATIVE).

In the SOME trials, weak cards involved nine symbols of the type that matched the [symbol] term. Strong and false cards contained nine symbols, three symbols of one type and six symbols of another type: on strong cards, the minority symbol type matched the [symbol] term whereas, on false cards, none of the symbols did. In the NUMBER trials, weak cards contained six symbols that matched the [symbol] term. Strong and false cards contained four symbols: on strong cards, these symbols matched the [symbol] term whereas, on false cards, they didn't. Finally, in the AD HOC trials, weak cards contained two different symbols, one of which matched the [symbol] term in the accompanying sentence. Strong and false cards contained a single symbol: on strong cards, this symbol matched the [symbol] term whereas, on false cards, it didn't. As a result, for each expression type, the symbol cards in the prime trials were configured in a similar fashion in the WEAK, STRONG and ALTERNATIVE prime trials in that they all involved one card in the strong card configuration and one card in the weak card configuration.

The design of the experiment involved two blocks of trials: a first block in which target covered card trials were unprimed (BASELINE conditions) and a second one in which these same covered card trials were primed. The rationale for this block design is that the target trials from Block 1 permit to measure how often quantity implicatures are observed without and prior to priming and thus to establish a baseline rate which can be then used to assess the direction and strength of the priming effects tested in Block 2.

Block 1 included only target and control trials. In the target trials, the overt card depicted a situation where the sentence presented is true only if it is read without a quantity implicature. Thus, the choice

[7] The priming potential of these two alternative types is further investigated in Experiment 2 where we explore what happens when the actual alternative to the target sentence is presented in the prime trials. In Experiment 1, we followed Rees and Bott (2018) in not using lexically identical versions of the alternative primes.

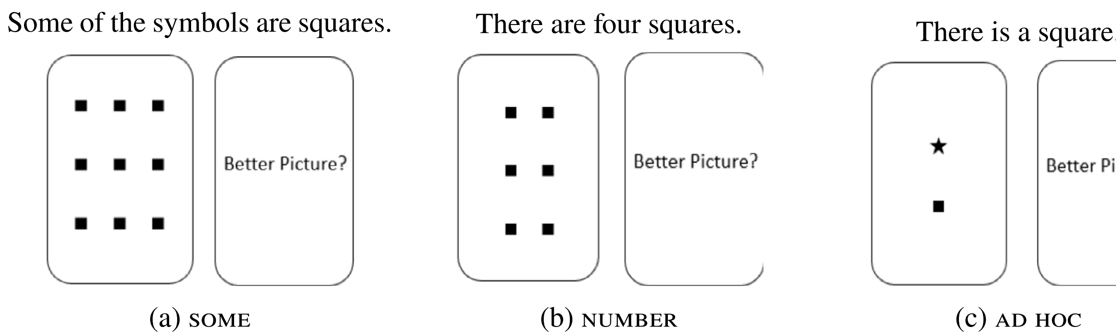

Fig. 3. Example target trials for SOME, NUMBER and AD HOC in the experiments reported in this paper. In these trials, the choice is between an overt card and a covered card. If participants interpret the sentence as conveying a quantity implicature, they should select the covered card; otherwise, they should select the overt card.

of the covered 'Better Picture?' card in these trials indicates that the participant has computed a quantity implicature. Block 1 contained four target trials for each of the three expressions – SOME, NUMBER, and AD HOC – corresponding to our BASELINE conditions (see Fig. 3). In addition, there were four true and four false control trials for each expression. In these trials, the overt card made the sentence either clearly true or clearly false, regardless of whether a quantity implicature was computed. Hence, there were 36 items in total in Block 1.

Block 2 involved target trials identical in structure to those in Block 1, but each target trial was preceded by two prime trials of the same kind. There were three kinds of prime trials: STRONG, WEAK, and ALTERNATIVE – all of which involved two overt pictures (see Fig. 4). STRONG and WEAK prime trials involved the same sentence frames as in the target trials. The pictures in these trials were constructed according to the following logic. In the STRONG prime trials, one of the overt cards made the sentence true without the quantity implicature of interest but false with it; the other card made the sentence true with or without the relevant implicature. Thus, in order to choose the latter card as the correct one, the participant has to compute the quantity implicature of interest, because otherwise both cards would make the sentence true. In the WEAK prime trials, the picture to be chosen had the same structure as the incorrect pictures in the STRONG primes and so it made the sentence true only if it is understood without a quantity implicature; the other picture simply made the sentence false with or without the relevant implicature, thereby forcing the participant to choose the former picture by suspending the quantity implicature of interest.

ALTERNATIVE primes involved alternative expressions as linguistic stimuli, together with two overt pictures, one of which made the sentence true while the other one made it false. For SOME, ALTERNATIVE prime trials involved a sentence of the form *All of the symbols are [symbol]*, which is more informative than the corresponding SOME sentence relative to the correct overt picture. Similarly, for NUMBER, ALTERNATIVE prime trials involved a sentence of the form *There are six [symbol]*, which is more informative than the corresponding NUMBER sentence relative to the correct overt picture. Finally, for AD HOC, as discussed, we created two kinds of ALTERNATIVE prime trials, X&Y-ALT and X-ALT. X&Y-ALT trials involved a conjunctive sentence of the form *There is a [symbol] and a [other symbol]* with two distinct symbol nouns. The sentence was presented with a 'true' card with two symbols matching the symbol types used in the sentence, and a 'false' card with a single non-matching symbol. X-ALT trials, on the other hand, involved a simpler sentence constructed by the same frame as the target AD HOC sentences, namely, *There is a [symbol]*. The sentence was presented with a 'true' card with a single matching symbol and a false card with two non-matching symbols.

STRONG and ALTERNATIVE priming conditions were tested separately from each other by adopting a partial between-subject design so that every participant saw the WEAK primes but only one of [1] the STRONG primes, [2] the ALTERNATIVE primes with X&Y-ALT for AD HOC, and [3] the ALTERNATIVE primes with X-ALT for AD HOC. In sum, each participant was presented with two prime types, WEAK and TEST, where TEST was one of Antoniou et al. (2016), Anwyl-Irvine, Massonnié, Flitton, Kirkham, and Evershed (2019) and Atlas and Levinson (1981) above. The rationale for manipulating the TEST primes between-subject, rather than within-subject, was the same as the one motivating our block design: it aimed to reduce the risk of uncontrolled, spillover effects between priming conditions and, specifically, the risk that participants' responses in the ALTERNATIVE conditions be affected by their exposure to STRONG prime stimuli in Block 2. Trials in Block 2 tested all three expressions in their WEAK and TEST priming conditions, with four iterations of each condition, giving rise to 72 triplets. Block 2 also contained filler trials. These trials were included to prevent participants from recognising the 'prime-prime-target' configuration of the experimental triplets. Filler trials were individual prime or target trials which were identical in all respects to the prime and target trials used in the priming conditions (see Figs. 3–5). Eight filler items were constructed for each expression, with half of them being individual prime trials (two WEAK and two TEST prime trials) and the other half being target trials. Thus, Block 2 consisted of 72 prime-prime-target triplets and 24 filler items.

For each trial, the symbol type used in the sentence was picked at random from our list of symbol types, with replacement across trials. The contents of the symbol cards accompanying each sentence were pseudo-randomly determined according to the relevant expression and the relevant condition: matching symbol types always corresponded to the symbol type used in the sentence while non-matching symbol types were randomly chosen from our list by excluding the matching symbol types. For each trial, the position of the two cards on the screen was chosen randomly.

5.2.4. Procedure

The experiment was run as an online survey using the Gorilla Experiment Builder (https://www.gorilla.sc; Anwyl-Irvine et al., 2019). The survey had two parts, one for each block of trials, with a self-timed break in between. Participants were given general instructions at the beginning of the survey and they were then given more specific instructions before starting each part (see Appendix A).

In the first part, participants were told that they would be presented with sentences, each of them would be accompanied by two pictures, one visible to them and another one covered with the text 'Better Picture?' on it. They were instructed to click on the visible picture if they considered it a match for the sentence, otherwise to click on the covered picture. Following these instructions, the experiment started with the trials from Block 1 (BASELINE conditions). In the second part, participants were told that, in some cases, both pictures would now be visible to them and they were instructed to click on the picture that they considered a better match for the sentence. Following these instructions, the experiment proceeded with the trials from Block 2 (WEAK and TEST priming conditions).

Participants were pseudo-randomly assigned one type of TEST prime so as to get an even number of participants for each of the three

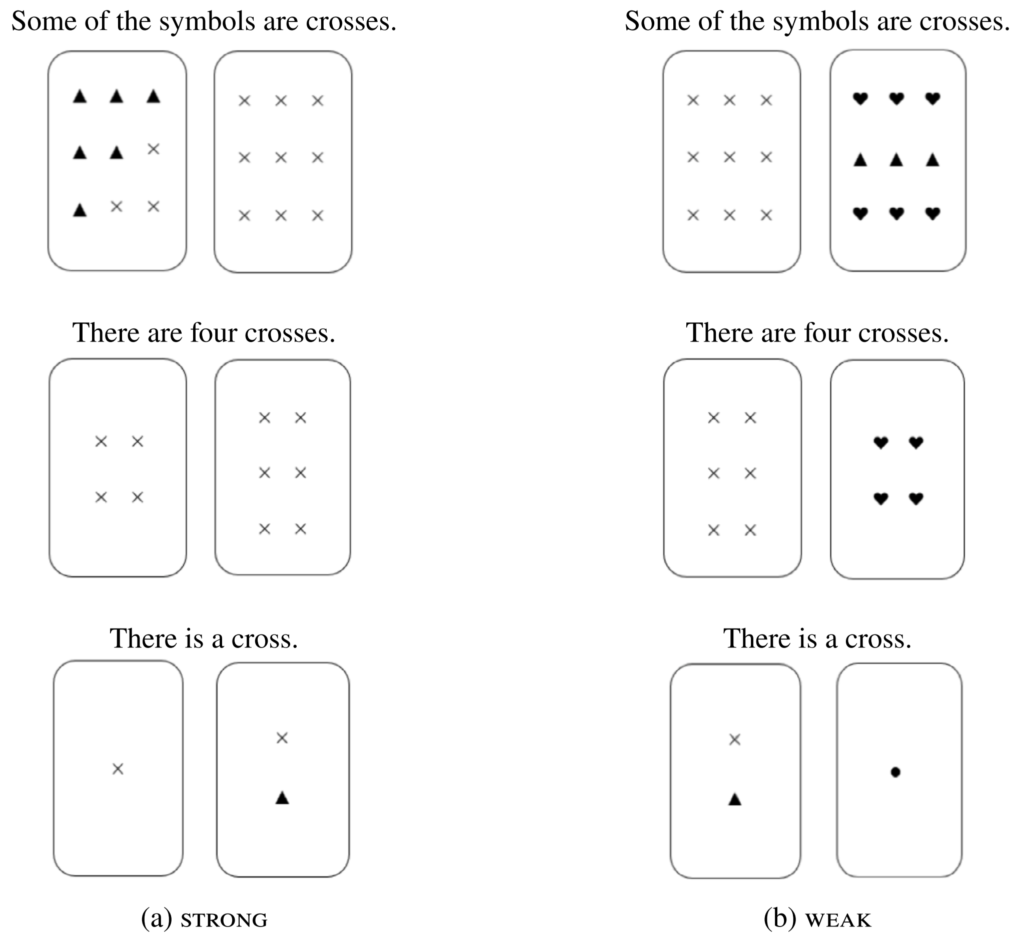

Fig. 4. Example (a) STRONG and (b) WEAK prime trials for SOME (top), NUMBER (middle), and AD HOC (bottom) in Experiment 1. In prime trials, the choice is between two overt cards. Participant choose the one that best fits the sentence. The expected choice here corresponds to the card on the left.

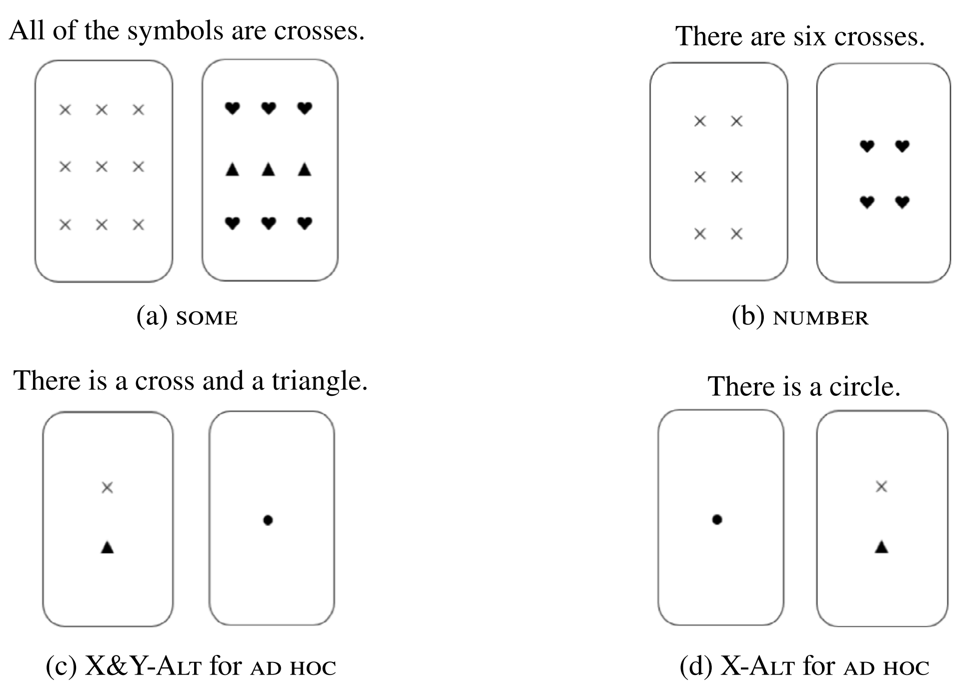

Fig. 5. Example ALTERNATIVE prime trials for (a) SOME, (b) NUMBER, (c) X&Y-ALT for AD HOC and (d) X-ALT for AD HOC in Experiment 1 (the expected choice corresponds here to the card on the left).

TEST primes (about 60 participants per TEST prime type). In each block, individual trials and triplets were presented in random order. On each trial, a fixation cross appeared and remained on the screen for 500 ms before the items were displayed. For each item, participants provided their response by clicking with the mouse on the picture of their choosing. Items remained on the screen until participants gave their response.

5.3. Results

5.3.1. Data treatment

Responses from 4 participant were excluded from analyses because their performance on the control trials from Block 1 did not reach the pre-established threshold of 80% accuracy. The mean accuracy rate of the remaining participants was above 98% for the True control trials and above 97% for the False control trials. Next, following the procedure discussed in Bott and Chemla (2016) and Raffray and Pickering (2010), we removed all responses to primed target trials that were not preceded by two correct prime responses. In total, 301 out of 4,200 responses to primed target trials were removed due to incorrect prime responses (which account for about 7% of the primed target trials, 5% of all target trials and 3% of the whole data set).

5.3.2. Data analyses

Analyses were conducted using the `R diptest` (Maechler, 2013), the `lme4` (Bates, Mächler, Bolker, & Walker, 2014; Bates, Maechler, & Bolker, 2011) and the `emmeans` (Lenth, Love, & Herve, 2017) libraries for the R statistics program (R. Core Team, 2021).

Based on the data analyses from pilot studies, we carried out preliminary tests for unimodality of the distribution of the by-participant mean rates in the BASELINE conditions by calculating the Hartigan dip-test statistic (Hartigan & Hartigan, 1985) .[8] For each of the three prime groups, the baseline rates in Block 1 were distributed unimodally for NUMBER (all Ds< 0.07, ns) and AD HOC (all Ds< 0.06, ns), but not for SOME (all Ds> 0.14, all ps< .001). Specifically, for SOME, there were two modes present in the BASELINE data, one peaking above 99% and the other below 1%. This is evidence that some participants consistently understood the SOME sentences with their quantity implicature, while others consistently understood them without it.[9] Thus, results to SOME trials were directly sorted according to the two responder profiles: participants were classified as Weak-Some responders if their baseline rate was below 50% and as Strong-Some responders if their baseline rate was above 50%. In total, there were 97 Strong-Some responders and 68 Weak-Some responders, roughly evenly distributed across all three groups, representing about 55% and 40% of the subjects in our sample, respectively. 10 participants had a baseline rate of exactly 50% (about 5% of the subjects in our sample), and so their results were not included in the analyses of the SOME target trials.

[8] Data and analysis code associated with the pilot studies we are referring to are available open access on the OSF platform at https://osf.io/263xf/.

[9] This finding is in line with the results of a number of studies reporting substantial variation in responses to *some*-sentences and, in some cases, a bimodal distribution between logical and pragmatic responders, both with adults (e.g., Guasti et al., 2005; Hunt, Politzer-Ahles, Gibson, Minai, & Fiorentino, 2013; Noveck & Posada, 2003) and children (e.g., Foppolo, Guasti, & Chierchia, 2012; Foppolo, Mazzaggio, Panzeri, & Surian, 2021; Guasti et al., 2005; Horowitz, Schneider, & Frank, 2018; Noveck, 2001). This finding, however, is far from systematic. As mentioned in Section 3, other studies have found that adult participants uniformly reject under-informative uses of 'some' (e.g., Papafragou & Musolino, 2003); similarly, results from inference tasks often show that adults readily endorse the 'not-all' implicatures associated with *some*-sentences, with little variation among participants (e.g., van Tiel et al., 2016). We submit that these variations in the distribution of by-participant rates are amenable to a similar task-based explanation to that we discussed in Section 3.

Table 2
Output of the model for NUMBER in Experiment 1. Condition and Prime group were coded with treatment contrasts using TEST as a reference level for Condition and ALTERNATIVE as a reference level for Prime group.

		Estimate	SE	z value	Pr(> |z|)
Number					
	(Intercept)	0.99	0.27	3.58	< .001
	Baseline	3.98	0.88	4.51	< .001
	Weak	−2.64	0.29	−8.93	< .001
	Prime group	1.13	0.49	2.31	< .05
	Prime group: Baseline	−1.64	0.83	−1.98	< .05
	Prime group: Weak	0.07	0.49	0.15	.88

Participants' responses were analysed by modelling response-type likelihood using logit mixed-effects regression models (Jaeger, 2008). Analyses primarily aimed at comparing – for each expression type and, in the case of SOME, for both responder profiles – the effect of the TEST primes relative to their corresponding BASELINE and WEAK conditions. For our purposes, the data for SOME and NUMBER from the two groups of ALTERNATIVE primes were aggregated as the ALTERNATIVE conditions for these expressions were identical across both groups. All models included Condition (3 levels: Baseline, Weak and Test), Prime group (2 levels for NUMBER and SOME: Strong, Alternative; 3 levels for AD-HOC: Strong, X&Y-Alt, X-alt) and their interaction as fixed factors, and the maximal random effect structure justified by the design and supported by the data, as recommended by Barr, Levy, Scheepers, and Tily (2013). For SOME, the maximal converging models included random intercepts for Subject; for NUMBER and AD HOC, the maximal converging models included random intercepts for Subject and random slopes for Condition grouped by Subject. Pairwise comparisons of the BASELINE, WEAK and TEST conditions were performed based on the estimated marginal means from the models we tested, and p-values were adjusted using the Bonferroni correction method for multiple testing.

5.3.3. NUMBER *trials*

Fig. 6 shows the proportion of 'Better Picture?' selection on the NUMBER trials by experimental condition and prime group. Model results are shown in Table 2.

Overall, the pattern of results for the STRONG and ALTERNATIVE prime groups were much alike. For both groups, the rates of 'Better Picture?' selection were the highest in the BASELINE conditions and the lowest in the WEAK conditions, with the TEST conditions somewhere in between. Yet both groups also differed in a critical way: the rates in the two priming conditions were more distant from the baseline rates in the ALTERNATIVE than in the STRONG group. Specifically, the model for NUMBER yielded a significant interaction between Condition and Prime group for the TEST vs. BASELINE comparison ($\beta = -1.64$, $SE = 0.83$, $p < .05$) showing that, relative to the relevant baselines, participants derived the quantity implicature for NUMBER sentences less often after ALTERNATIVE than after STRONG primes. On the other hand, there was no such an interaction for the TEST vs. WEAK comparison ($\beta = 0.07$, $SE = 0.49$, ns), suggesting that the strength of the contrasts between WEAK and TEST conditions – that is, of the main priming effects – was essentially similar in both groups.

This interpretation of the model results is further supported by the results of the pairwise comparisons that we carried out between the three levels of the Condition factor. In both groups, the rates of 'Better Picture?' selection in the WEAK conditions were significantly lower than in the TEST conditions (all $|\beta|$s> 2.56, all adjusted ps< .001), establishing the presence of priming effects, and significantly lower than in the BASELINE conditions (all $|\beta|$s> 4.9, all adjusted ps< .001), showing that the relevant effects are driven by the WEAK primes. Furthermore, the difference between BASELINE and TEST conditions was significant in the ALTERNATIVE group ($\beta = -3.99$, SE = 0.88, adjusted $p < .001$) and marginally significant in the STRONG group ($\beta = -2.34$, SE= 0.99, adjusted $p = .054$), thus evidencing the presence of spillover effects from the WEAK primes onto the TEST conditions as well.

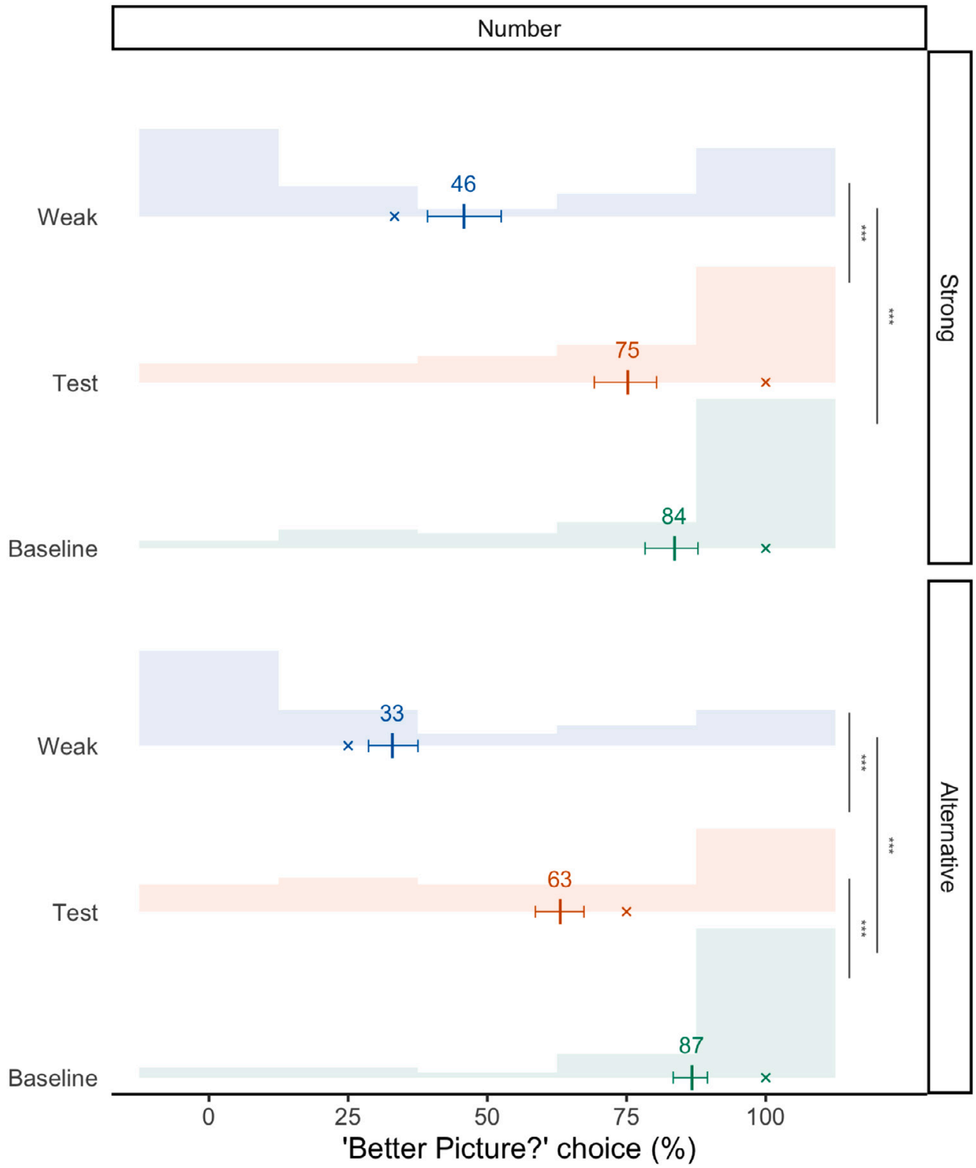

Fig. 6. Proportion of 'Better Picture?' selection on NUMBER trials in Experiment 1 by condition (BASELINE, WEAK, TEST) and prime group (STRONG, ALTERNATIVE). For each condition, the distribution of by-participant mean proportions is visualised by a histogram, the grand mean by a thick bar with its rounded value on top and the 95% CI around it, and the median by a cross. The significance levels are based on the adjusted p-values for all pairwise comparisons tested in each prime group.

These results establish that, for NUMBER, the various priming effects observed in both prime groups are driven by participants' exposure to the WEAK primes and that the WEAK primes had wider-ranging effects in promoting the initially less preferred, weak interpretation of NUMBER sentences across priming conditions. Returning to the questions raised leading into our study, and assuming that baseline responses here reflect prior expectations or biases towards strong interpretations for NUMBER sentences, we can say that the direction of priming is driven by a shift in bias after encountering WEAK prime trials. Crucially, our results also show that covered-card choices on target trials which follow ALTERNATIVE primes and STRONG primes are significantly lower than the BASELINE, indicating a spillover effect of the WEAK prime stimuli on target trials beyond those that immediately follow. Moreover, ALTERNATIVE and STRONG primes differ in their ability to prompt the quantity implicature associated with NUMBER sentences, with the former being less effective than the latter at counteracting the spillover effects induced by the WEAK primes in the course of the experiment.

5.3.4. SOME *trials*

Recall that, for the SOME trials, participants were sorted into two responder profiles, Weak-Some and Strong-Some responders, based on their baseline preferences, i.e., their preferred interpretation prior to being exposed to prime trials. Fig. 7 shows the proportion of 'Better Picture?' selection on the SOME trials by experimental condition, prime group and responder profile. Model results are shown in Table 3.

For each responder profile, the patterns of results for the STRONG and ALTERNATIVE groups were very much alike, yet with some noticeable differences between the two, pointing here again to a difference in strength between ALTERNATIVE and STRONG primes. Starting with the Strong-Some responders (Fig. 7, right panel), the results for these participants were entirely parallel to those we found for NUMBER (see 5.3.3). The model yielded a significant interaction between Condition and Prime group for the TEST vs. BASELINE comparison ($\beta = -2$, $SE = 0.56$, $p < .001$), but not for the TEST vs. WEAK comparison ($\beta = -0.2$, $SE = 0.46$, ns). In both prime groups, the estimated means of 'Better Picture?' selection were significantly lower in the WEAK conditions than in the

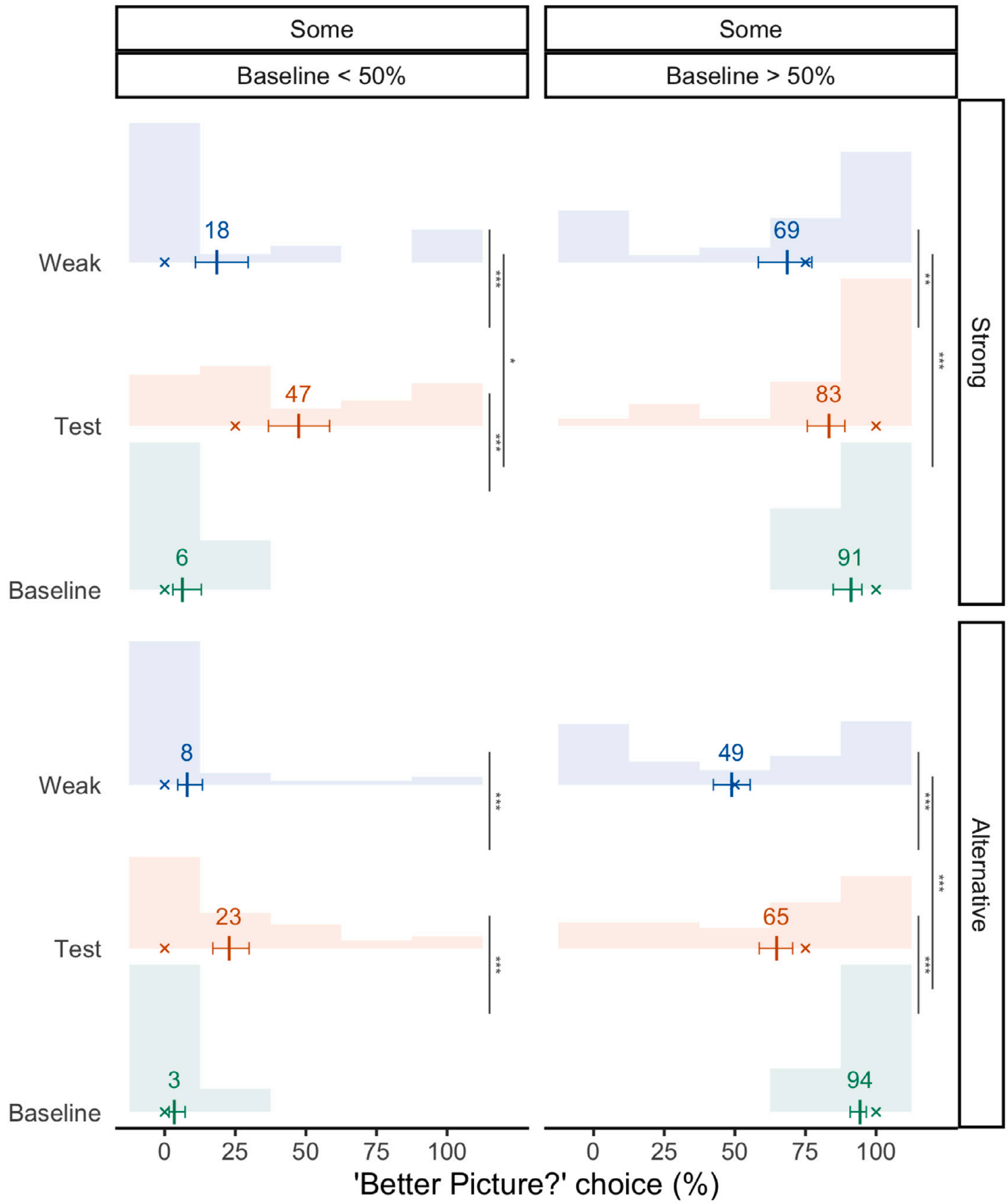

Fig. 7. Proportion of 'Better Picture?' selection on some target trials in Experiment 1 by condition, prime group and responder profile. This graph reads in an analogous way to the previous one (see Fig. 6 for details).

test (all $|\beta|$s> 0.98, all adjusted ps< .01) and baseline conditions (all $|\beta|$s> 2.1, all adjusted ps< .001). The difference between baseline and test conditions was significant in the alternative group ($\beta = -2.92$, $SE = 0.35$, adjusted $p < .001$), but not in the strong group ($\beta = -0.92$, $SE = 0.44$, ns). In sum, these results show that, for the Strong-Some responders, the weak primes gave rise to below-baseline rates of 'Better Picture?' selection in both prime groups and substantially affected responses to the test conditions in the alternative group. These findings provide further evidence that alternative primes were less efficient than strong primes at counteracting the spillover effects induced by the weak primes.

Turning to the Weak-Some responders (Fig. 7, left panel), the results for these participants were the mirror image of those found for the Strong-Some responders both in the strong and in the alternative group. The model didn't yield any significant interaction between Condition and Prime group, but it yielded a main effect of Prime group ($\beta = 1.38$, $p < .05$) which, in the absence of any interaction, can be interpreted meaningfully. Specifically, this effect shows that, in contrast to the Strong-Some speakers, Weak-Some responders selected significantly more often the covered card in the strong than in the alternative group. In both prime groups, the estimated means of 'Better Picture?' selection were significantly higher in the strong conditions than in the weak (all $|\beta|$s> 1.59, all adjusted ps< .001) and baseline conditions (all $|\beta|$s> 2.71, all adjusted ps< .001), evidencing the presence of priming effects driven by the strong primes and thus in the opposite direction to those observed for number and for the Strong-Some responders. Finally, the difference between baseline and weak conditions was significant in the strong group ($\beta = 1.57$, SE= 0.59, adjusted $p < .05$), but not in the alternative group ($\beta = 1.12$, SE= 0.56, ns). Taken together, the outcomes for the Strong-Some and Weak-Some responders support the view that

Table 3
Output of the model for SOME in Experiment 1. Both factors were dummy coded. For Condition, WEAK was used as a reference level for the Weak-Some responders while TEST was used for the Strong-Some responders. For Prime group, ALTERNATIVE was used as a reference level for both responder profiles.

		Estimate	S.E.	z value	Pr(> $|z|$)
Weak-Some					
	(Intercept)	−3.37	0.47	−7.05	< .001
	Baseline	−1.12	0.56	−1.98	< .05
	Test	1.59	0.41	3.83	< .001
	Prime group	1.38	0.67	2.05	< .05
	Prime group: Baseline	−0.44	0.81	−0.54	.58
	Prime group: Test	0.18	0.63	0.29	.76
Strong-Some					
	(Intercept)	1.02	0.28	3.62	< .001
	Baseline	2.92	0.35	8.28	< .001
	Weak	−0.98	0.24	−3.97	< .001
	Prime group	1.22	0.51	2.36	< .05
	Prime group: Baseline	−2.00	0.56	−3.54	< .001
	Prime group: Weak	−0.20	0.46	−0.43	.66

Table 4
Output of the model for AD HOC in Experiment 1. Both factors were coded with treatment contrasts using WEAK as a reference level for Condition and X&Y-ALT for Prime Group.

		β	S.E.	Z	p-value
Ad-hoc					
	(Intercept)	−6.94	1.12	−6.14	< 001
	Baseline	−1.74	1.60	−1.09	.27
	Test	2.95	0.99	2.95	< .01
	Strong	2.68	1.08	2.46	< .05
	X-Alt	1.13	1.04	1.09	.27
	Strong: Baseline	−2.99	1.60	−1.86	.06
	Strong: Test	−1.01	0.75	−1.34	.18
	X-Alt: Baseline	−0.80	1.55	−0.51	.6
	X-Alt: Test	−2.04	0.73	−2.76	< .01

prime stimuli have a larger impact on covered card choice when they force participants to attribute readings to sentences containing *some* contrary to what their prior bias is. This is the same outcome that we found in the NUMBER condition. In the priming literature, such outcomes are referred to as *inverse preference effects* (a.o., Hartsuiker & Kolk, 1998; Hartsuiker & Westenberg, 2000; Scheepers, 2003). In addition, the results from the Weak-Some responders provide some evidence that ALTERNATIVE primes are also less effective than STRONG primes at prompting the quantity implicature associated with SOME sentences.

5.3.5. AD HOC *trials*

Fig. 8 shows the proportion of 'Better Picture?' selection on the AD HOC trials by experimental condition and prime group (STRONG, X&Y-ALT and X-ALT). Model results are shown in Table 4.

The model for AD HOC yielded a significant interaction between Condition and Prime group for X&Y-ALT vs. X-ALT groups in the WEAK vs. TEST conditions ($\beta = -2.04$, SE= 0.73 $p < .01$), showing that the difference between WEAK and TEST conditions was larger in the X&Y-ALT than in the X-ALT group. The interaction between Condition and Prime group was also marginally significant for X&Y-ALT vs. STRONG groups in the WEAK vs. BASELINE conditions ($\beta = -2.99$, SE= 1.60, $p = .06$), suggesting that the difference between WEAK and BASELINE conditions was larger in the STRONG than in the X&Y-ALT group. No other significant or marginally significant interactions between Condition and Prime group were found.

In line with the model results, the relationships between all three conditions were found to be remarkably different in each prime group. For the STRONG group, all pairwise comparisons were significant (BASELINE<WEAK<STRONG, all $|\beta|$s> 1.93, all adjusted ps< .05) so that the STRONG primes boosted the rates 'Better Picture?' selection above baselines across priming conditions, in a way similar to what we found for SOME among the Weak-Some responders. For the X&Y-ALT group, only the pairwise comparisons involving the TEST conditions came out as significant ($|\beta|$s> 2.95, adjusted ps< .01). These results show that the X&Y-ALT primes had a moderate, yet detectable boosting effect; however, in contrast to what we found in the STRONG group, there is no evidence in our data that the boosting effect of the X&Y-ALT primes affected participants' behaviour beyond the TEST conditions (i.e., no evidence of spillover effects). Finally, for the X-ALT group, none of the pairwise comparisons reached significance.[10] Thus, there is no evidence in our data that X-ALT primes affected in any remarkable way participants' behaviour in the priming conditions.

Finally, we would like to touch on another interesting aspect of the present data which may be of interest for future work. As pointed out to us by an anonymous reviewer, the rates of implicature derivation for AD HOC in the target trials, here as well as in Experiments 2 and 3 below, are remarkably low compared to those observed in acquisition studies like Stiller, Goodman, and Frank (2011, 2015), Horowitz et al. (2018) and Foppolo et al. (2021), where *ad hoc* implicatures are found instead to be readily computed not only by adults, but also by children at a young age. We note, however, that all the studies mentioned above are based on picture selection tasks in which, in the target trials, a weak and a strong picture are presented side-by-side to the participants, usually with an additional false picture used as a distractor. Therefore, the analogue of these trials in our and previous priming studies are not the baseline (or primed) target trials but instead the STRONG prime trials. Crucially, our results for these prime trials fully align with those reported in the above studies in showing that, in such cases, participants systematically favoured the strong over the weak symbol card, with a selection rate above 95%. Thus, what the present data actually shows is that the rates of 'weak card' selection were *overwhelmingly* higher in the target than in the STRONG prime trials, suggesting that the use of a covered card in place of a strong one had a substantial effect on participants' responses. These findings may teach us that, in the absence of a picture supporting the stronger reading, subjects less often entertained the type of alternatives involved in *ad hoc* implicature or, similarly, the type of context making these implicatures relevant (see Section 8 for further discussion of this point).

5.4. Discussion

The present study reproduces in full the priming effects from previous research in showing that, for all three expressions that we investigated, participants consistently provided more responses based on quantity implicatures after STRONG and canonical ALTERNATIVE primes than after WEAK primes. As we explained, however, the objective of this study lay somewhere else: it aimed to elucidate the direction of priming in these differences, by the use of a baseline, which was itself not susceptible to any potential spillover effects. Our findings in this regard shed new light on the direction and strength of previously observed priming effects. In addition, we have evidence for spillover effects within the second block of trials. In this discussion, we will unpack our findings a little further and discuss how we should amend, in significant ways, the interpretation of these effects proposed in the models of Bott and Chemla (2016) and Rees and Bott (2018).

First, let us consider the comparison between WEAK and STRONG primes. It seems clear that differences between these conditions are essentially *inverse preference effects*. The baseline rate for NUMBER sentences was quite high, suggesting that participants generally had low expectations that such sentences have a weak reading. Upon encountering trials where they were forced to endorse the weak reading of these sentences,

[10] For completeness, we note that the contrast between BASELINE and TEST in that group was marginally significant (p = .07 after correction). While we cannot rule out the possibility that this contrast be of empirical interest, we also note that it cannot be given too much importance in the context of our study in the absence of evidence for priming or spillover effects in the X-ALT data.

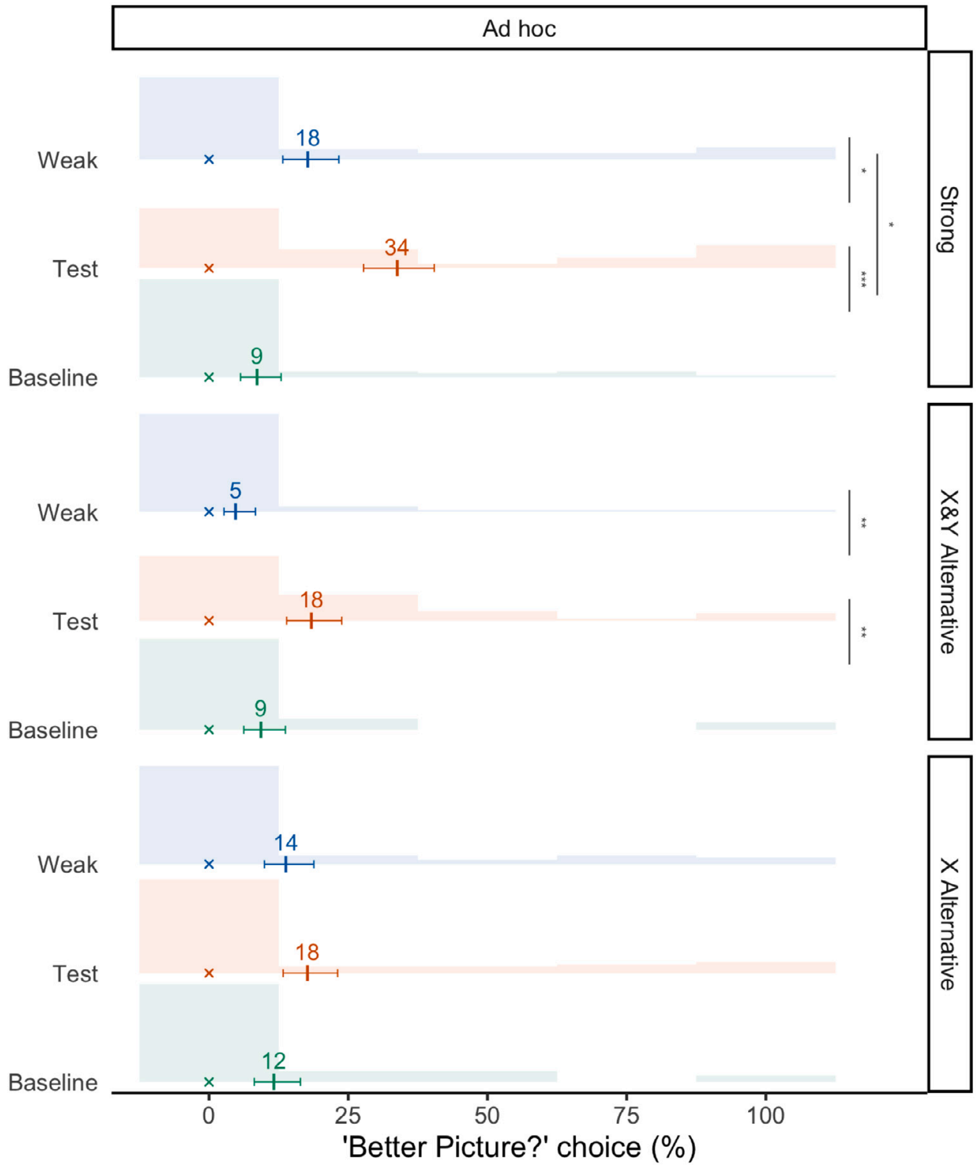

Fig. 8. Proportion of 'Better Picture?' selection on AD HOC target trials in Experiment 1 by condition and prime group. This graph reads in an analogous way to the previous ones (see Fig. 6 for details).

participants' expectations (in the context of the experimental session) were shifted so that they were more willing to choose the open card both after WEAK prime trials and, to a lesser extent, even after STRONG prime trials. Likewise, for those participants whose initial bias was to interpret SOME sentences with implicature, encountering WEAK prime trials raises expectations that lead to without-implicature readings for the stimuli sentences. By contrast, in AD HOC trials, the baseline rate was quite low, indicating that target AD HOC stimuli biased our participants to a weak reading. In this case, it is the prime stimuli forcing the reading with a quantity implicature that shifted expectations away from the baseline, leading to more covered card choices after STRONG primes. Likewise, for those participants whose initial bias was to interpret SOME sentences without implicature, it is the prime trials forcing the dispreferred reading that lead to the largest shift in behaviour.

Overall, given the findings in our baseline block of trials, we can conclude that previously observed differences between STRONG and WEAK priming conditions manifest priming in both directions. In particular, where the baseline is high (NUMBER and, for some responders, SOME), we see large priming effects driven by exposure to WEAK primes as well as spillover effects after other prime trials resulting from the presence of WEAK primes in the second block. This outcome is challenging for the Salience Hypothesis, which relies on the salience of linguistic alternatives as the main causal mechanism of priming. In addition, we note that, if we were to explain this outcome by supposing that inverse priming effects are the result of priming non-salience of alternatives, then it would mean that alternatives should generally be regarded as non-salient in normal speech situations. This, however, runs contrary to the fact that we often do derive quantity implicature with little discourse context. Apart from the salience of alternatives, Bott and Chemla (2016) proposed that a common mechanism for quantity implicature, namely the mechanism of negation, can also be primed. However, by their own estimation, priming of this mechanism accounts for only a small proportion of their within-scale priming effects. What we observe in Experiment 1 is that within-scale priming in the WEAK condition for NUMBER is very strong, and accounts for the entire STRONG–WEAK priming effect. Thus, neither the salience of alternatives nor the activation/deactivation of the mechanism of exclusion provide a satisfactory basis for understanding these priming effects.

The overall inverse-preference effect observed for STRONG and WEAK primes can find an account in the Context Adaptation Hypothesis outlined above. Simply put, results in our BASELINE condition for the three expression types demonstrate that AD HOC sentences and, for around half of our participants, SOME sentences are strongly associated with contexts in which alternative propositions are not relevant. On the other hand, NUMBER sentences and, for around half of our participants, SOME sentences are strongly associated with contexts in which alternative propositions are relevant and thus excluded, leading to a more informative update. WEAK primes require stimuli sentences to be understood in contexts of the first kind, while STRONG primes require contexts of the second. Where the baseline bias is already for weak contexts, then only STRONG primes are likely to have a discernible effect on shifting expectations. Where the baseline bias is for strong contexts, only WEAK primes may shift expectations.

Another important finding of this study, which is replicated in both experiments below, is that, in trials that followed priming towards the baseline-dominant reading of the sentence, covered card choices were different from the baseline rate for the NUMBER and AD HOC conditions. NUMBER trials after STRONG primes elicited significantly fewer covered card choices than BASELINE, while AD HOC trials after WEAK primes elicited significantly more covered card choices than BASELINE. These outcomes, which we described as 'spillover effects', hardly make sense when priming effects are considered in terms of salience. However, they can be easily explained when we adopt a context-adaptation approach: in the second block of trials, the adjustment of expectations prompted by the less preferred reading affects not only the trials immediately following the prime in question, but all similar trials across the whole session.

Next, turning to the results of the two groups where the test stimulus followed ALTERNATIVE primes, the first thing we can say is that our results extend the findings from the STRONG condition in showing that alternative priming effects obey the same logic as implicature priming effects: just like the WEAK-STRONG contrasts, the WEAK-ALTERNATIVE contrasts are driven in this experimental paradigm by the prime type promoting speakers' less preferred interpretation prior to being exposed to prime trials. Thus, for NUMBER sentences, the WEAK primes were found to drive the priming effects of interest by promoting the initially less preferred, weak interpretation of these sentences. On the other hand, for AD HOC sentences, the with-implicature interpretation was generally the less preferred interpretation prior to priming and, for these sentences, the STRONG and X&Y-ALT ALTERNATIVE primes were found to be the driving force behind the contrasts we observed. Finally, for SOME, both kinds of priming effects were found to co-exist in our data, with a principled distribution across the two responder profiles we identified (Strong-Some vs. Weak-Some). These findings disconfirm the assumption in Rees and Bott (2018) that ALTERNATIVE primes, simply by virtue of making stronger alternatives salient, systematically lead to boosting implicature derivation in the following target trials. Conversely, these findings support the view that, just like implicature priming effects, alternative priming effects are inverse preference effects, the direction of which depends on speakers' prior preferences.

Our results also show that ALTERNATIVE and STRONG primes tend to differ in their ability to prompt implicature computation. This is evidenced in our data by the fact that the rates of 'Better Picture?' selection in the priming conditions were generally lower in the ALTERNATIVE than in the STRONG prime group. In the case of the TEST-driven priming effects (i.e., SOME for the Weak-Some responders and AD HOC), this finding indicates that ALTERNATIVE primes were generally less effective at priming the relevant quantity implicatures; in the case of the WEAK-driven priming effects (i.e., SOME for the Strong-Some responders and NUMBER), this finding indicates that ALTERNATIVE primes were also less effective at counteracting the spillover effects from the WEAK primes. These observations show us two sides of the same coin: compared to the baselines, ALTERNATIVE primes induced smaller changes in behaviour than the STRONG primes. Again, these findings go against Rees and Bott's proposal that alternative salience is all that is needed to explain priming effects, and, more generally, against the idea that ALTERNATIVE and STRONG primes would have identical priming effects. They suggest instead that some distinction must be maintained between, on the one hand, prior processing of alternative sentences and, on the other, prior processing of sentences which involves implicature computation. On a context-adaptation approach, this distinction can be captured, by assuming that, compared to the WEAK and STRONG primes, ALTERNATIVE primes provide speakers with less informative cues as to the type of conversational context they're in, resulting in weaker adaptation effects. For example, in the case of SOME, whereas STRONG primes are associated with QuDs such as *what proportion of symbols are squares* and WEAK primes with QuDs such as *whether some of the symbols are squares*, ALTERNATIVE primes are a little more equivocal about whether the QuD is about the proportion of symbols, as in STRONG primes, or about a specific proposition, as in WEAK primes. In the latter case, what may be carried over to the target trial is a specific question, *whether all symbols* are of a certain type, or a more general yes/no question schema.

Finally, the results for AD HOC, as well as SOME for the Weak-Some responders, provide some positive evidence that alternative prime trials can prompt quantity reasoning and facilitate meaning enrichment. Based on our findings, however, this phenomenon appears to be observed only under specific conditions: the weak interpretation of the sentence of interest must be the dominant interpretation prior to ALTERNATIVE prime trials. The fact that ALTERNATIVE primes did not have the same boosting effect across the board speaks against the Salience Hypothesis and, subsequently, against the characterisation of alternative priming effects proposed in Rees and Bott (2018), which was based on that hypothesis. However, before reaching the conclusion that alternative salience is not at all a factor in these priming effects, there is an important consideration that should be taken into account.

In our study, just like in Rees and Bott's experiments, the symbol types used in the prime and target trials were randomly chosen so that the symbols used in the prime trials were generally different from those used in the target trials. Thus, in the general case, the sentences presented in the ALTERNATIVE prime trials were not lexically identical to the alternatives that participants actually had to generate in order to derive the quantity implicatures of interest in the following target trials. In this regard, the fact that X&Y-ALT primes were found to give rise to alternative-driven priming effects can be taken to show that identity of content words between the prime and the (alternative of the) target sentence is not always a prerequisite for such effects to arise. This, however, does not rule out the possibility that lexical identity is a factor that matters. In particular, it could be a prerequisite for the logically independent X alternatives to AD HOC that we tested, because the ability of these alternatives to enter quantity reasoning entirely depends on their lexical content. This, in turn, could explain why we were not able to find any priming effect associated with X-ALT primes in the present experiment. The aim of Experiment 2 was to determine if there may be alternative priming effects over and above those observed in Experiment 1 if we ensure that alternative prime stimuli are lexically identical to the actual alternatives involved in the derivation of the implicature associated with the target sentence.

We turn finally to that aspect of our study which compares X&Y with X alternative sentences, where we see a priming effect for the former but not for the latter. It could be argued that the difference observed does provide some evidence that salience of alternative is a factor in implicature priming effects. The reasoning would be that, *pace* Fox and Katzir (2011), only more informative alternatives can serve as alternatives for *ad hoc* implicatures. In that case, the difference could be attributed to the X&Y-ALT primes having some effect activating the relevant alternative for the target trial. We note that this line of argument cannot explain the challenges which inverse preference effects pose to the Salience Hypothesis more generally. But we do accept that, in the particular case of *ad hoc* implicatures, it is possible that salience is a factor, in addition to context adaptation, due to the fact that *ad hoc* implicatures generally rely on determining a particular set of alternatives in context. This is to be compared with the other two scalar types which, it could be argued, have pre-existing strong associations with their alternatives. While this line of reasoning deserves to be explored further, we believe that the effects in question can find a ready explanation on the Context Adaptation Hypothesis: the conjunctive sentences involved in the X&Y-ALT primes generally gave rise to more *ad hoc* implicatures than the simpler sentences involved in the X-ALT primes because these sentences are more readily perceived as an attempt of the speaker to exhaustively list all relevant symbols. If conjunctive sentences indeed have a higher association with QuDs that make *ad hoc* implicatures relevant (compared to non-conjunctive sentences), then the X&Y-ALT primes may not have worked as genuine alternative primes in our experiment, but rather as some kind of STRONG primes in disguise, by directly priming the kind of QuDs favouring meaning strengthening. This line of explanation, therefore, would account for the presence of above-baseline priming effects with X&Y-ALT primes and for the absence of such effects with X-ALT primes independently of alternative-related considerations. In Experiment 3, we tested these competing hypotheses about the role of X&Y alternatives in promoting SI responses.

6. Experiment 2: Investigating the role of lexical identity

While the results from Experiment 1 cast doubt on the Salience Hypothesis, they do not rule out the possibility that alternative salience is an independent factor in some way, the effect of which could be revealed only when the actual alternatives to the target sentence are used in the prime trials. The goal of Experiment 2 was to refine the results from the ALTERNATIVE conditions of Experiment 1 and further test the impact of alternative salience on implicature computation by investigating the potential effect of lexical identity on alternative priming. We wanted to know whether alternative primes are more effective when they involve sentences that are lexically identical to the alternatives involved in the derivation of the implicature of the target sentence.

For these purposes, Experiment 2 retested the BASELINE, WEAK and ALTERNATIVE conditions from Experiment 1 while manipulating the lexical correspondence between the symbol types appearing in the prime and in the target trials of the ALTERNATIVE conditions. Concretely, instead of being randomly chosen as in Experiment 1, symbol types in the ALTERNATIVE conditions were manipulated so as to create two types of prime-prime-target triplets. In the DIFFERENT triplets, the symbol nouns in the ALTERNATIVE primes were systematically different across prime trials and different from the symbol type required to generate the appropriate alternative in the following target trials. On the other hand, in the SAME triplets, the symbol nouns in the ALTERNATIVE primes were the same across prime trials and the same as the symbol type relevant to generate the alternative of interest. We reasoned that, if alternative salience boosts the derivation of quantity implicatures, then lexical identity should make ALTERNATIVE primes more effective and, consequently, the SAME triplets should show a greater priming potential than the DIFFERENT triplets either directly, by further boosting the rates of implicature derivation, or indirectly, by further counteracting the spillover effects coming from the WEAK primes.

6.1. Data availability

Stimuli, data, and analysis code for Experiment 2 are all freely available on the OSF platform at https://osf.io/6gsv9/.

6.2. Methods

6.2.1. Participants

192 novel participants (120 female, average age 36 years) were recruited online through Prolific using the same pre-screening criteria as in Experiment 1. Participants were paid £1.40, and average completion time was about 9 min. The consent and data collection procedures were the same as in Experiment 1.

6.2.2. Materials and design

Experiment 2 was based on the same materials and block design as Experiment 1 (see 5.2.3 for details). In particular, the composition of the two blocks of trials in this novel experiment was identical to that of Block 1 and Block 2 in Experiment 1.

Block 2 involved two kinds of prime trials, WEAK and ALTERNATIVE, both of which were constructed in an analogous fashion as those from Experiment 1. As before, there were two kinds of ALTERNATIVE primes for AD HOC, X&Y-ALT and X-ALT, which were tested separately from each other, as in Experiment 1. That is, each participant was presented with two prime types, WEAK and TEST, where TEST corresponded to ALTERNATIVE primes with either X&Y-ALT or X-ALT for AD HOC. Crucially, in contrast to Experiment 1, the symbol types in the ALTERNATIVE conditions were selected in a principled way so as to manipulate the identity of the content words used in the linguistic stimuli of the prime and target trials.

There were two types of ALTERNATIVE triplets, DIFFERENT and SAME. In the DIFFERENT triplets, the symbol nouns were systematically different across prime trials and different from the symbol type relevant to generate the appropriate alternative in the following target trials. Thus for instance, for the X-ALT primes of AD HOC, a DIFFERENT triplet could look as follows: *There is a star* (Prime 1) → *There is a cross* (Prime 2) → *There is a circle* (Target), where the target sentence was presented with an overt card depicting a circle and another symbol type distinct from those already mentioned in the primes, e.g., a square. By contrast, in the SAME triplets, the symbol nouns were the same across prime trials and the same as the symbol type relevant to generate the alternative of interest. Thus, for the X-ALT primes of AD HOC, a SAME triplet could look as follows: *There is a square* (Prime 1) → *There is a square* (Prime 2) → *There is a circle* (Target), where the target sentence was presented with an overt card depicting a circle and another symbol type matching the one mentioned in both primes, i.e., a square. DIFFERENT and SAME triplets were created for all three expressions and, in the case of AD HOC, for the two alternative types of interest. Just like the type of TEST primes, the type of ALTERNATIVE triplets presented in the experiment (SAME vs. DIFFERENT) was manipulated between-subject so as to reduce the risk of uncontrolled, spillover effects between priming conditions. The rest of the design was identical in all relevant respects to that of Experiment 1.

6.2.3. Procedure

At the beginning of the study, participants were pseudo-randomly assigned one type of TEST prime (ALTERNATIVE primes with either X&Y-ALT or X-ALT for AD HOC) and one type of TRIPLET (either DIFFERENT or SAME) so as to get an even number of participants for each of the four possible combinations (about 48 participants per combination). The rest of the procedure was identical to the one used in Experiment 1 (see Section 5.2.4 and Appendix A for the instructions).

6.3. Results

6.3.1. Data treatment

Data treatment was the same as for Experiment 1 (see 5.3.1 for details). 1 participant was excluded due to low performance on the control trials from Block 1 (accuracy < 80%). The mean accuracy rate of the remaining participants was above 98% for the True control trials and above 96% for the False control trials. 237 out of 4,584 responses to primes target trials were removed due to incorrect prime responses (about 5% of the primed target trials, 3.5% of all target trials and 2% of the whole data set).

6.3.2. Data analyses

Data analyses were essentially the same as in Experiment 1 (see 5.3.2 for details). In line with the data from Experiment 1, the by-participant baseline rates for SOME in this experiment were not distributed unimodally (all Ds > 0.15, all ps $< .001$). Thus, once again,

Table 5
Output of the model for NUMBER in Experiment 2. Condition and Triplet type were coded with treatment contrasts using TEST as a reference level for Condition and DIFFERENT as a reference level for Triplet.

		Estimate	SE	z value	Pr(> $\|z\|$)
Number					
	(Intercept)	−0.33	0.34	−0.968	.33
	Baseline	3.02	0.28	10.78	< .001
	Weak	−1.30	0.22	−5.71	< .001
	Triplet	0.97	0.48	2.00	< .05
	Triplet: Baseline	−0.29	0.38	−0.76	.44
	Triplet: Weak	−0.62	0.33	−1.86	.06

results of SOME trials were sorted according to two responder profiles, Strong-Some and Weak-Some responders, exactly as before. In total, there were 91 Strong-Some responders and 82 Weak-Some responders, roughly evenly distributed across all four groups of conditions, representing about 47% and 43% of the subjects in our sample, respectively.[11]

As in the data analyses of Experiment 1, the data for SOME and NUMBER from the two groups of ALTERNATIVE primes were aggregated as the ALTERNATIVE conditions for these two expressions were identical in all respects across both groups. The SOME data were analysed by distinguishing Strong-Some and Weak-Some responders, and the AD HOC data by distinguishing X&Y-ALT primes and X-ALT primes. All models included Condition (3 levels: Baseline, Weak and Test), Triplet type (2 levels: Different, Same) and their interaction as fixed factors, and a random intercept for Subject.[12] The procedure for performing pairwise comparisons between BASELINE, WEAK and TEST conditions and for correcting p-values was the same as in Experiment 1.

6.3.3. NUMBER *trials*

Fig. 9 shows the proportion of 'Better Picture?' selection on the NUMBER trials by condition and triplet type. Model results are shown in Table 5.

The patterns of results for the DIFFERENT and the SAME triplets were very similar to one another and to the pattern of results observed for NUMBER in the ALTERNATIVE priming conditions of Experiment 1. No significant interaction between Condition and Triplet type was found for either of the two-way comparisons involving the TEST conditions, suggesting that the Triplet manipulation did not affect in any remarkable way the size of the contrasts between the different levels of the Condition factor. Thus, there is no evidence that, for NUMBER, participants' responses in the ALTERNATIVE conditions were affected by the type of triplets they were presented with, nor is there evidence that one type of triplets was more effective than the other at counteracting the spillover effects from the WEAK primes onto the ALTERNATIVE conditions.

For both types of triplets, the rates of 'Better Picture?' selection were significantly higher in the BASELINE than the TEST and WEAK conditions (all $|\beta|$s< 2.73, adjusted ps< .001), and significantly higher in the TEST than WEAK conditions (all $|\beta|$s> 1.31, adjusted ps< .001). These results are in line with those from Experiment 1 in showing that (i) participants readily derived the quantity implicatures associated with NUMBER sentences prior to priming, (ii) the WEAK-ALTERNATIVE contrast for these sentences was driven by the WEAK primes, and (iii) the most effective prime type, here the WEAK primes, gave rise to wider-ranging, spillover effects.

[11] 9 participants tested on the DIFFERENT triplets and 9 participants tested on the SAME triplets had a baseline rate of exactly 50%. The responses of these participants (about 9% of the subjects in our sample) were not included in the analyses of the SOME target trials.

[12] This corresponded to the maximal random effect structure supported by the data.

Table 6
Output of the model for SOME in Experiment 2. Both factors were dummy coded. For Condition, WEAK was used as a reference level for the Weak-Some responders while TEST was used for the Strong-Some responders. For Triplet, DIFFERENT was used as a reference level for both responder profiles.

		Estimate	S.E.	z value	Pr(> $\|z\|$)
Weak-Some					
	(Intercept)	−2.74	0.44	−6.14	< .001
	Baseline	−0.89	0.44	−1.99	< .05
	Test	1.15	0.37	3.07	< .01
	Triplet	−0.17	0.60	−0.28	.77
	Triplet: Baseline	−0.24	0.67	−0.37	.71
	Triplet: Test	0.49	0.53	0.91	.36
Strong-Some					
	(Intercept)	0.66	0.35	1.84	.06
	Baseline	3.67	0.50	7.24	< .001
	Weak	−1.61	0.32	−4.96	< .001
	Triplet	0.70	0.50	1.41	.15
	Triplet: Baseline	−0.93	0.64	−1.46	.14
	Triplet: Weak	0.01	0.44	0.03	.97

6.3.4. SOME *trials*

Fig. 10 shows the proportion of 'Better Picture?' selection on the SOME trials by condition, triplet type and responder profile. Model results for the Weak-Some and the Strong-Some responders are shown in Table 6.

For each responder profile, the patterns of results for the DIFFERENT and SAME triplets were much alike and much like those observed in the ALTERNATIVE priming conditions of Experiment 1. The model results were also similar for both responder groups in showing significant contrasts between the different levels of the Condition factor, but no significant interaction between Condition and Triplet. Thus, the SOME data offer no evidence that participants' responses in the ALTERNATIVE conditions were affected by the type of triplets they were presented with in these conditions.

As in Experiment 1, the results for the Strong-Some responders (Fig. 10, left panel) were entirely parallel to those observed for NUMBER: for both types of triplets, the rates of 'Better Picture?' selection were significantly higher in the BASELINE than in the TEST and WEAK conditions (all $|\beta|$s> 2.74, adjusted ps< .001), and significantly higher in the TEST than in the WEAK conditions (all $|\beta|$s> 1.60, adjusted ps< .001), evidencing the presence of priming and spillover effects driven by the WEAK primes. Once again, the direction of these effects was reversed for the Weak-Some responders (Fig. 10, right panel): the rates of 'Better Picture?' selection were significantly higher in the TEST than in the WEAK and BASELINE conditions for both types of triplets (all $|\beta|$s> 1.15, adjusted ps< .01) and marginally higher in the WEAK than in the BASELINE conditions for the SAME triplets (β = 1.14, adjusted p = .06), but not for the DIFFERENT triplets (β = 0.89, adjusted p = .13). In sum, these results confirm the results for SOME in Experiment 1 and offer further experimental support in favour of the view that alternative priming effects are inverse preference effects.

6.3.5. AD HOC *trials*

Fig. 11 shows the proportion of 'Better Picture?' selection on the AD HOC trials by condition, triplet type and alternative type. Model results for the X&Y and X alternatives are shown in Table 7.

The rates for AD HOC were comparable to those observed in Experiment 1 in showing that participants generally preferred the interpretation of AD HOC sentences without quantity implicatures, independent of the condition, type of alternatives and type of triplets they were presented with. This general impression is confirmed by the model results. Neither the model for X&Y-ALT, nor the one for X-ALT yielded any significant interaction between Condition and Triplet type for either of the two-way comparisons involving the TEST conditions, i.e., the ALTERNATIVE primes. Thus, in line with the results for NUMBER and SOME, the manipulation of the Triplet factor had no effect on participants' responses to AD HOC trials either.

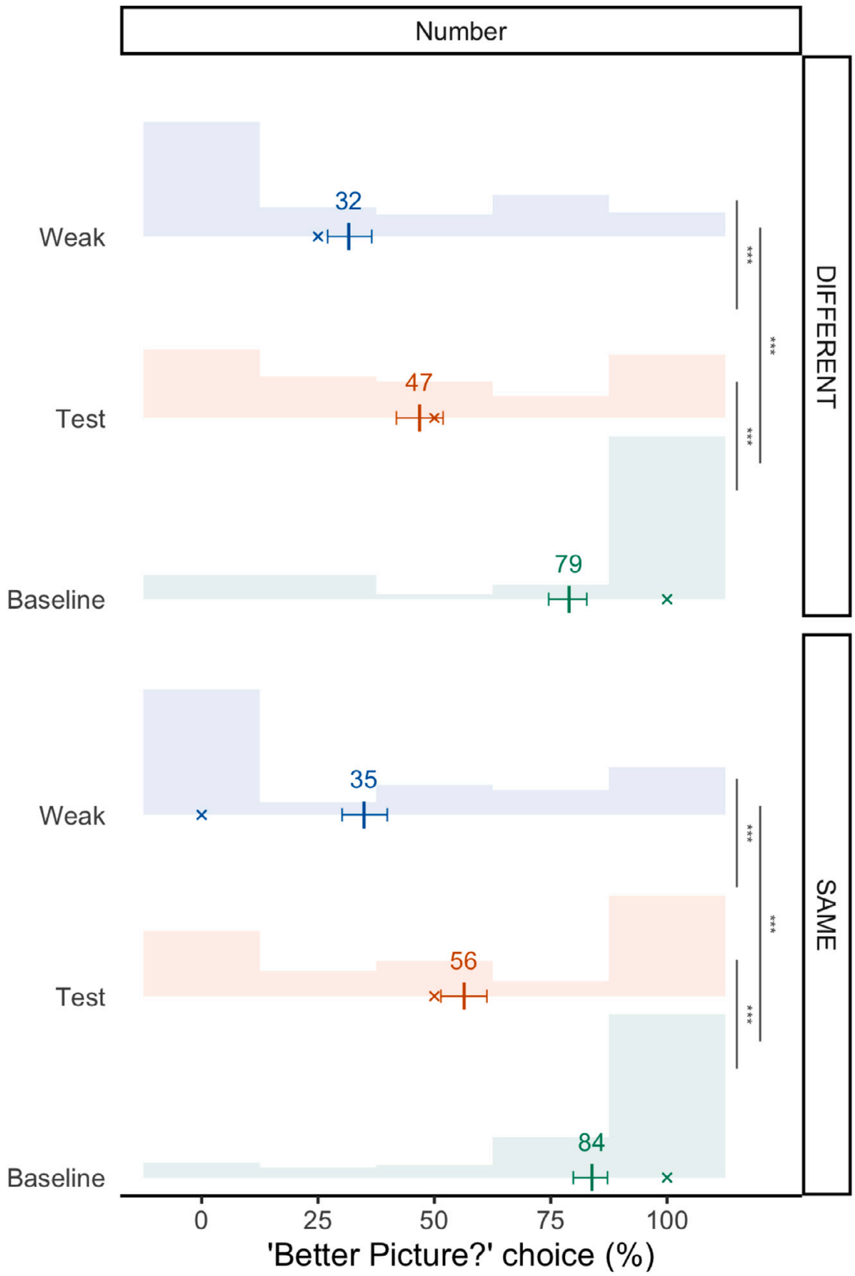

Fig. 9. Proportion of 'Better Picture?' selection on NUMBER trials in Experiment 2 by condition and triplet type. This graph reads in an analogous way to the previous ones (see Fig. 6 for details).

The pairwise comparisons yielded similar results as in Experiment 1. For the X&Y alternatives, the rates of 'Better Picture?' selection were significantly higher in the TEST than in the WEAK and BASELINE conditions for both types of triplets (all $|\beta|$s> 1.25, all adjusted ps< .05), establishing the presence of ALTERNATIVE-driven priming effects. For the X alternatives, the rates of 'Better Picture?' selection were significantly higher in the TEST than in the WEAK conditions for the DIFFERENT triplets ($\beta = 1.93$, $p < .001$), and marginally so for the SAME triplets ($\beta = 1.02$, $p = .06$). For the DIFFERENT triplets, this contrast demonstrates a genuine priming effect, but one driven by the WEAK primes, which gave rise to below-baseline rates ($\beta = -1.68$, $p < .01$). For the SAME triplets, however, the present data do not allow us to determine a particular direction for the observed contrast as no reliable difference was found between priming and baseline conditions (WEAK vs. BASELINE: $\beta = -0.21$, adjusted $p = 1$; TEST vs. BASELINE: $\beta = 0.80$, adjusted $p = .18$).

6.4. Discussion

These results reproduce the main findings from Experiment 1 and suggest that lexical identity has no significant impact on the effectiveness of ALTERNATIVE primes. Specifically, we found that ALTERNATIVE primes involving sentences lexically identical to the alternatives of the target sentences (SAME triplets) had a similar priming potential as those that did not (DIFFERENT triplets), even in cases where the most preferred reading prior to the priming phase was the weak reading. These findings add strength to the conclusion that alternative salience has no effect on the derivation of quantity implicatures in this paradigm. Since lexical identity does not improve the effectiveness of alternative primes in any remarkable way, we conclude that variations in the lexical contents of the prime sentences cannot be the reason for not finding more robust priming effects for ALTERNATIVE primes in general, nor can it be the reason

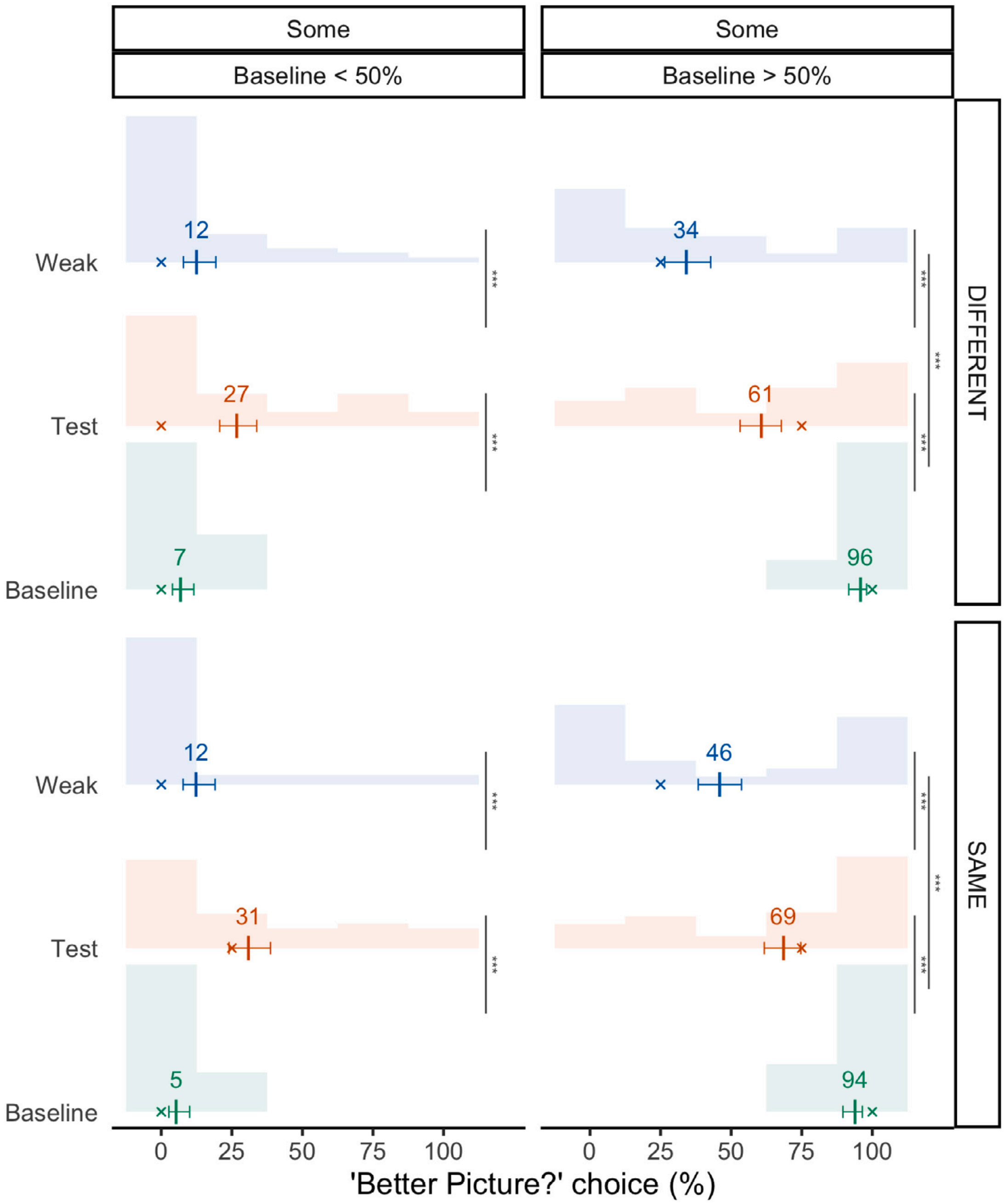

Fig. 10. Proportion of 'Better Picture?' selection on SOME target trials in Experiment 2 by condition, triplet type and responder profile. This graph reads in an analogous way to the previous ones (see Fig. 6 for details).

for the contrasts that we observed between the X-ALT and X&Y-ALT priming conditions of AD HOC.

Before moving on, however, let us anticipate and address a possible concern regarding the interpretation of the AD HOC data in this experiment. The concern comes from the fact that, for the SAME triplets, the present data is suggestive of a contrast between WEAK and X-ALT primes and, at the same time, inconclusive as to the direction of this potential contrast. As such, this data leaves open the possibility that X-ALT primes had in fact a small boosting effect, which we simply failed to detect in our experiment, e.g., because our study did not have enough power to detect such small effects. To address this concern, we partially rerun Experiment 2 by retesting the X-ALT primes for AD-HOC with 200 participants, doubling the subject sample size for both triplet types compared to the original study. The materials, design and procedure were thus identical to those used in the original study except for the X&Y-ALT primes, which we did not retest. As in the original study, there was no significant interaction between Condition and Triplet type for either of the two-way comparisons of interest (all $|\beta|$s< 0.89, ns). Most importantly, for the SAME triplets, the rate of 'Better Picture?' selection was significantly lower in the WEAK conditions than in both the X-ALT and the BASELINE conditions (adjusted ps$< .01$), thus revealing a priming effect driven by the WEAK primes. This additional data offers conclusive evidence that, irrespective of their lexical contents, the X-ALT primes failed to induce above-baseline priming effects.[13]

[13] The data was analysed in a similar way as the data from the original study. The results for NUMBER and SOME were parallel to those found in the original study and are thus left aside in the interest of space. The data and

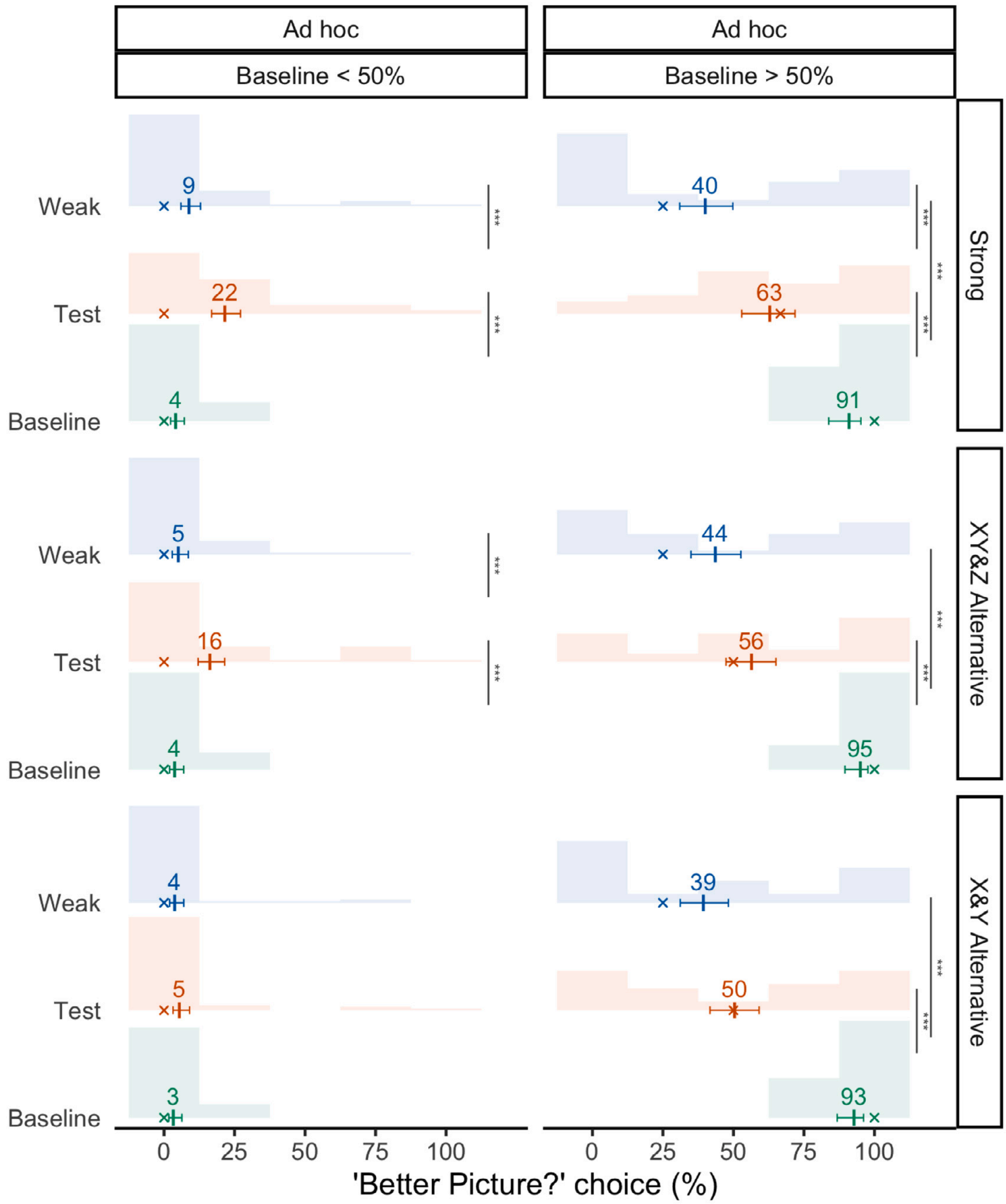

Fig. 11. Proportion of 'Better Picture?' selection on AD HOC target trials in Experiment 2 by condition, triplet type and alternative type. This graph reads in an analogous way to the previous ones (see Fig. 6 for details).

As previously discussed, there is a further question about why X&Y-ALT primes have a greater effect than X-ALT primes. Assuming, for the sake of argument, that logically independent sentences cannot serve as genuine alternatives for ad hoc implicatures, it may be that the observed difference in priming effect is due to some small contribution of salience after all. Specifically, salience of the conjunctive form may have increased the salience of the right kind of alternative. Our counter-proposal, based on the Context Adaptation hypothesis is that X&Y-ALT primes were more effective at eliciting inferences about context that leads to the implicature in the target. We test the predictions of these competing hypotheses in the next study.

analysis script for this experiment can be found at the same address as before (https://osf.io/6gsv9/).

7. Experiment 3: controlling for the effect of conjunction

The goal of Experiment 3 was to explore further the AD HOC results from Experiment 1 to see if X&Y-ALT prime stimuli were effective in raising implicature rates in the target by eliciting the right kind of context, or raising the conjunctive form to salience. For these purposes, Experiment 3 was designed by minimally modifying the AD HOC trials from Experiment 1 so that prime and target AD HOC trials all involved conjunctive sentences. The novel sentences were obtained by adding a conjunct to the prime and target AD HOC sentences tested in Experiment 1. Thus, the novel target sentences for AD HOC were conjunctive sentences similar to those used in the X&Y-ALT primes of Experiment 1, as exemplified in (4), while the novel prime sentences were (i) X&Y alternatives structurally identical to the target sentences and (ii) XY&Z alternatives involving three conjuncts and therefore structurally more

Table 7
Output of the model for AD HOC in Experiment 2. Both factors were coded with treatment contrasts using TEST as a reference level for Condition and DIFFERENT for Triplet.

		Estimate	S.E.	z value	Pr(> |z|)
X&Y Alternative					
	(Intercept)	−5.01	1.19	−4.19	< .001
	Baseline	−2.64	0.63	−4.13	< .001
	Weak	−1.64	0.56	−2.92	< .01
	Triplet	0.51	0.99	0.51	.60
	Triplet: Baseline	1.38	0.76	1.81	.06
	Triplet: Weak	−0.39	0.74	−0.53	.59
X Alternative					
	(Intercept)	−3.59	0.66	−5.38	< .001
	Baseline	−0.25	0.38	−0.65	.51
	Weak	−1.93	0.50	−3.84	< .001
	Triplet	−0.03	0.74	−0.04	.96
	Triplet: Baseline	−0.55	0.58	−0.95	.33
	Triplet: Weak	0.91	0.67	1.36	.17

complex than the target sentences, as exemplified in (4a) and (4b), respectively.

(4) There is a cross and a triangle. AD HOC
⇝ *There is a cross, a triangle and nothing else*

a. There is a heart and a circle. X&Y-ALT
b. There is a heart, a circle and a diamond. XY&Z-ALT

As in Experiment 1, target sentences appeared in the baseline block as well as after prime trials in the second block. According to the Context Adaptation Hypothesis, *there*-sentences which contain conjunction (X&Y-ALT) are more suggestive of an attempt to be exhaustive than simple *there*-sentences (X-ALT). Thus, we should expect to see raised rates of implicature response in the baseline trials, compared to when simple *there*-sentences were used in Experiments 1 and 2. According to the Salience Hypothesis, we should see the same low rates in the baseline block as in the previous experiments, since in all cases, no alternatives are particularly salient at the baseline phase. As for the second block, our hypothesis predicts that there may be a small boost after XY&Z primes, over and above baseline. This is due to the fact that having a third conjunct is arguably a stronger cue to the right kind of exhaustive context, given the visual content of the stimuli (see Fig. 12). The alternative hypothesis being considered here is based on the idea that previous priming effects in AD HOC trials was a result of prime trials making a more informative alternative available. Thus, the prediction is similar for both hypotheses that the more complex conjunction in Block 2 will have a greater priming effect.

7.1. Data availability

Stimuli, data, and analysis code for Experiment 3 are all freely available on the OSF platform at https://osf.io/6gsv9/.

7.2. Methods

7.2.1. Participants

Anticipating that the ALTERNATIVE conditions for AD HOC may give rise to relatively weak priming effects (if any effect at all), as the results of Experiment 1 and 2 suggest, we increased the number of participants in this study. 299 novel participants (146 female, average age 40 years) were recruited online through Prolific using the same pre-screening criteria as in Experiment 1 and 2. Participants were paid £1.40, and average completion time was about 9 min. The consent and data collection procedures were the same as in Experiment 1 and 2.

7.2.2. Materials and design

Experiment 3 was designed by reproducing all the conditions for all three expressions from Experiment 1 while adjusting the contents of the AD HOC trials to our present purposes. Example target and prime trials for AD HOC are given in Fig. 12.

The novel AD HOC sentences were conjunctive sentences of the form *There is a [symbol] and a [other symbol]* with two distinct symbol nouns. In the target trials, the overt card depicted three different symbols, two of which were matching symbols, thus making these sentences true only if they are read without a quantity implicature. In the control trials, the overt card depicted either two different matching symbols, making these sentences clearly true, or two different non-matching symbols, making them clearly false. The STRONG and WEAK primes were constructed by the same logic as before using the AD HOC sentence frame and a combination of the overt cards involved in the target and controls trials: a 'true' and a 'false' card in the WEAK primes, and a 'true' and a 'target' card in the STRONG primes (see Figs. 12(a) and 12(b)). The ALTERNATIVE primes, namely X&Y-ALT and XY&Z-ALT, were designed on the basis of the X-ALT and X&Y-ALT primes from Experiment 1. X&Y-ALT primes involved a conjunctive sentence constructed by the same frame as the AD HOC sentences. The sentence was presented with a 'true' card with two different matching symbols and a 'false' card with three different non-matching symbols (see Fig. 12(c)). XY&Z-ALT primes involved more complex conjunctive sentences of the form *There is a [symbol], a [other symbol] and a [other symbol]* with three distinct symbol nouns. The sentence was presented with a 'true' card with three different matching symbols and a 'false' card with two different non-matching symbols (see Fig. 12(d)). Hence, all AD HOC prime trials involved one card with two symbols and one card with three symbols so that the cards in these trials were configured in a parallel fashion across all prime types, exactly as in previous experiments.

The rest of the design was identical in all respects to that of Experiment 1. Thus, the description of the design of Experiment 1 also stands as a description of that of Experiment 3, except for the modifications of the contents of the AD HOC prime and target trials that we just described. We refer the reader to Section 5.2.3 for details about the composition of the two blocks of trials, the distribution of the TEST priming conditions between subjects, and the pseudo-randomisation methods used to determine the contents and position of the cards.

7.2.3. Procedure

The procedure was identical to the one used in Experiment 1 (see Section 5.2.4 for details and Appendix A for the instructions).

7.3. Results

7.3.1. Data treatment

Data treatment was the same as for Experiment 1 and 2 (see Section 5.3.1 for details). 5 participants were excluded prior to analyses due to low performance on the control trials from Block 1 (accuracy $< 80\%$). The mean accuracy rate of the remaining participants was above 98% for the True control trials and above 97% for the False control trials. 450 out of 7,056 responses to primes target trials were removed due to incorrect prime responses (about 6% of the primed target trials, 4% of all target trials and 2.5% of the whole data set).

7.3.2. Data analyses

Data analyses were the same as in Experiment 1 (see Section 5.3.2 for details). As before, the by-participant baseline rates for SOME showed a non-unimodal distribution (all Ds> 0.10, all ps< .001) with 172 Strong-Some responders and 105 Weak-Some responders, representing about 58% and 35% of the subjects in our sample, respectively. In contrast to what we found in Experiment 1 and 2, however, the by-participant baseline rates for AD HOC in this experiment also showed a non-unimodal distribution (all Ds> 0.08, all ps< .001) with one peak above 99% and another peak below 1%. Accordingly, the results for

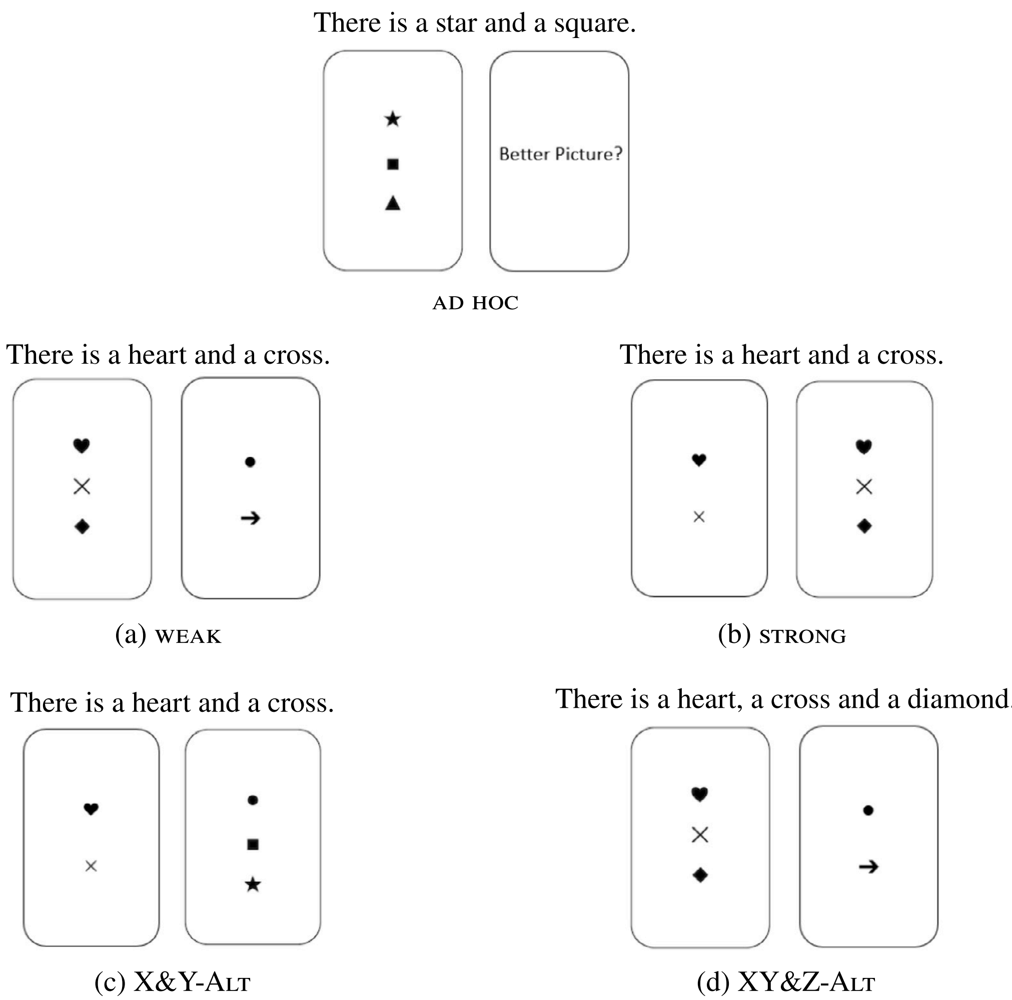

Fig. 12. Example target trial (top) and example prime trials for AD HOC in the (a) WEAK, (b) STRONG, (c) X&Y-ALT and (d) XY&Z-ALT conditions of Experiment 3.

AD HOC were also sorted into two responder profiles, Strong-Adhoc and Weak-Adhoc, according to the criteria previously established (baseline > 50% vs. baseline < 50%). In total, there were 86 Strong-Adhoc and 188 Weak-Adhoc responders, roughly evenly distributed across all three TEST primes, representing 29% and 63% of the subjects in our sample, respectively.[14]

In the following, we report on the results of the conditions of primary interest, namely the novel AD HOC conditions that were introduced in this experiment. The analyses of the NUMBER and SOME trials are provided in full in the analysis script associated with this experiment. The results for NUMBER and SOME were similar to those found in Experiment 1 and 2, replicating all the main findings reported so far for these trials and supporting the same conclusions as before.

7.4. AD HOC *trials*

Fig. 13 shows the proportion of 'Better Picture?' selection on the AD HOC trials by experimental condition, prime group and responder profile. Model results for the Weak and the Strong responders are shown in Table 8.

Overall, the baseline rates for AD-HOC sentences in this experiment were higher than those observed for the simpler AD-HOC sentences in Experiment 1 and 2 (33% on average contra 10% on average in Exp. 1–2). In relation to this first observation is the finding mentioned above that, in contrast to Experiment 1 and 2, the by-participant baseline rates for AD-HOC in this experiment were bimodally distributed with a substantial group of Strong-Adhoc responders showing baseline rates above 90% (see left panel in Table 8). Taken at face value, these observations suggest that, in the absence of any form of priming, conjunctive sentences are more likely to be interpreted with their AD-HOC implicatures than simpler, non-conjunctive ones. This result confirms our context-adaptation hypothesis about priming effects and challenges accounts based on salience of alternatives.

Turning now to the priming conditions, the response patterns for the Strong-Adhoc responders were very similar to those observed for the Strong-Some responders in this and our previous experiments: for all three prime types, the rates of 'Better Picture?' selection in the WEAK and TEST priming conditions were below the baseline rates, suggesting here again the presence of spillover effects from the WEAK primes onto the TEST priming conditions. For the Weak-Adhoc responders, on the other hand, above-baseline priming effects were found for the STRONG and XY&Z-ALT primes, but not for the X&Y-ALT primes. Thus, of the two conjunctive prime types we tested, only the more informative XY&X-ALT primes were found to modify speakers' prior preferences and boost the derivation of *ad hoc* implicatures.

7.5. Discussion

The results of this experiment are twofold. First, we found that the baseline rates for AD-HOC sentences in this experiment were much higher

[14] 20 participants had a baseline rate of exactly 50% (about 7% of the subjects); their responses were not included in the analyses of the AD HOC target trials that we report on below.

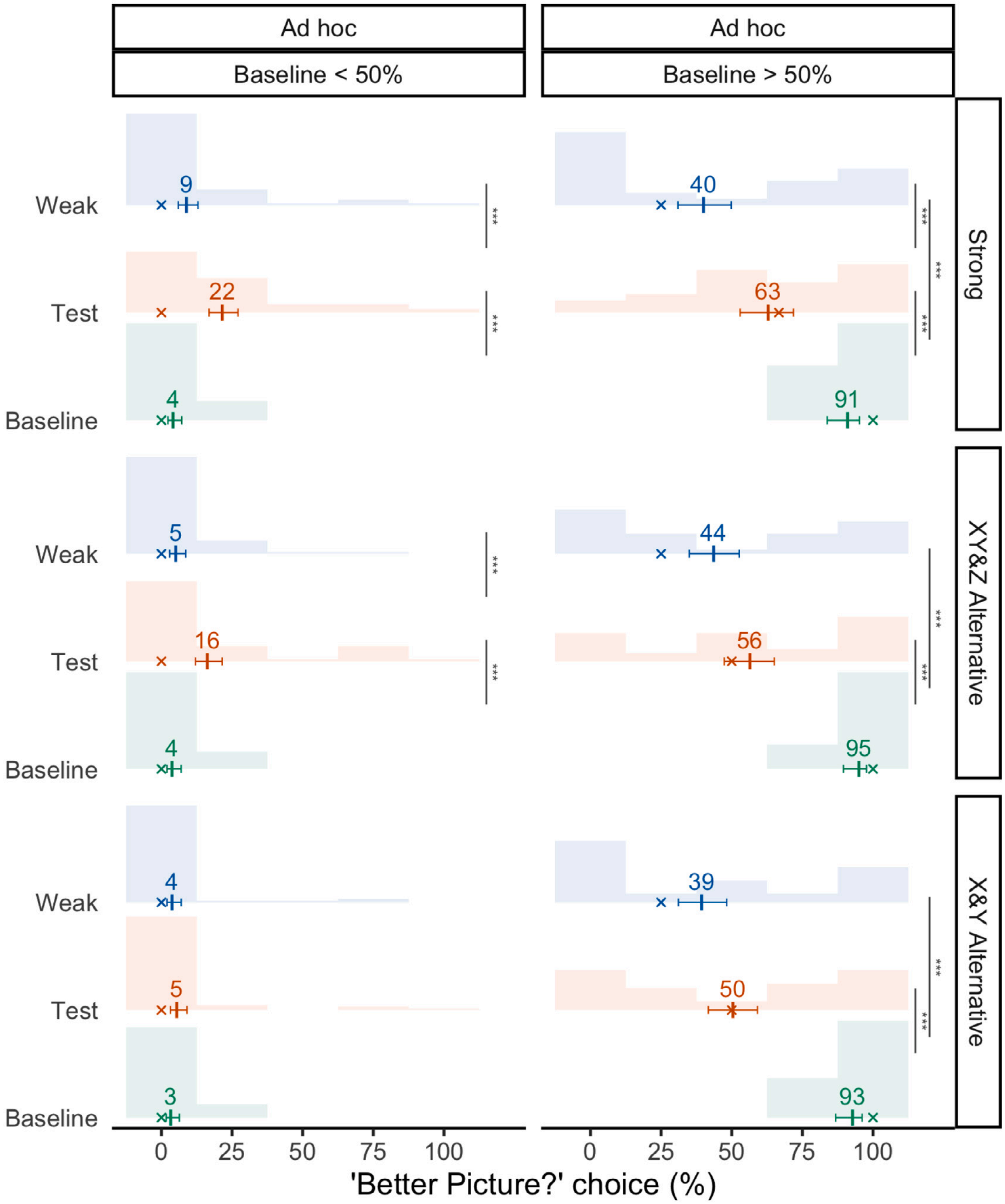

Fig. 13. Proportion of 'Better Picture?' selection on AD HOC target trials in Experiment 3 by condition, prime group and responder profile. This graph reads in an analogous way to the previous ones (see Fig. 7 for details).

than those observed in Exp. 1–2, indicating that X&Y sentences were more readily interpreted exhaustively than X sentences. This finding supports the idea that, compared to non-conjunctive forms, conjunctive forms have a higher prior on contexts that require exhaustivity. Second, despite the increased likelihood of conjunctive sentences to be interpreted exhaustively, we found that, in cases where the baseline rates were low (Weak-Adhoc responders), XY&Z-ALT primes had a noticeable boosting effect on participants' implicature rates while X&Y-ALT primes did not. These results better support the context-adaptation hypothesis, since it seems clear that conjunction itself can considerably increase participants propensity to derive implicatures, without the benefit of priming.

8. General discussion

Let us first summarise the main findings that we argue are problematic for the Salience Hypothesis. Putting aside for now the results of different types of alternative priming for *ad hoc* implicatures, the results of Experiment 1 show, thanks to the novel baseline conditions, that alternative priming generally gives rise to an inverse preference pattern, and this was replicated in the other two experiments. This observation already poses an issue for the Salience Hypothesis. Firstly, contrary to what is expected under the Salience Hypothesis, not only Strong and Alternative primes, but also Weak primes can have robust priming effects, when the baseline rate is high, and their priming effects are in the opposite direction, towards the interpretation without

Table 8
Output of the model for AD HOC in Experiment 3. Both factors were dummy coded. For Condition, WEAK was used as a reference level for the Weak responders while TEST was used for the Strong responders. For Prime group, XY&Z-ALT was used as a reference level for both responder profiles.

		Estimate	S.E.	z value	Pr(> |z|)
Weak-Adhoc					
	(Intercept)	−4.187	0.46	−9.02	< .001
	Baseline	−0.34	0.48	−0.71	.47
	Test	1.65	0.39	4.25	< .001
	STRONG	0.88	0.54	1.61	.10
	X&Y-ALT	−0.59	0.64	−0.91	.36
	STRONG: Baseline	−0.57	0.63	−0.91	.36
	X&Y-ALT: Baseline	0.19	0.74	0.25	.79
	STRONG: Test	−0.39	0.49	−0.79	.42
	X&Y-ALT: Test	−1.14	0.64	−1.79	.073
Strong-Adhoc					
	(Intercept)	4.19	0.62	6.71	< .001
	Weak	−4.44	0.56	−7.87	< .001
	Test	−3.58	0.54	−6.60	< .001
	STRONG prime	−0.81	0.82	−0.96	.33
	X&Y-ALT prime	−0.20	0.83	−0.24	.80
	STRONG prime: Weak	0.49	0.75	0.65	.51
	X&Y-ALT prime: Weak	−0.22	.77	−0.29	.76
	STRONG prime: Test	1.19	0.71	1.66	.09
	X&Y-ALT prime: Test	−0.27	0.74	−0.37	.70

quantity implicature. Such inhibition effects are hard to explicate under the Salience Hypothesis, and some other mechanism would have to be postulated. Secondly, when the baseline rate is high, we actually do not observe priming effects of Strong and Alternative priming, which is also contrary to what is expected under the Salience Hypothesis. In fact, in such cases, we observed the opposite of the prediction, that is, a rate of implicature responses lower than the baseline rate. We attributed this to spillover effects, which are priming effects of preceding Weak prime trials that have lingering effects. While the presence of spillover effects themselves is in principle compatible with the Salience Hypothesis, it needs to be acknowledged that, in order to maintain the Salience Hypothesis, one would have to assume that the effect of salience must be much weaker than spillover effects triggered by what happened prior to the presentation of the alternative in the prime trials and can be overridden by them. Importantly, this entails that increased salience of an alternative is not sufficient to kick-start the computation of a quantity implicature, which is contrary to the autonomous role of alternative salience in the generation of quantity implicatures that Rees and Bott (2018) seem to envisage.

In addition, the fact that alternative priming was observed in our results to be generally less effective than implicature priming poses further issues for the Salience Hypothesis. That is, Strong prime trials do not verbally present the relevant alternative itself, but simply force the interpretation with the target quantity implicature, the generation of which, by hypothesis, involves reference to the alternative. On the other hand, in Alternative prime trials, the alternative is verbally presented, so it is naturally expected that the Alternative prime makes the alternative more salient than the Strong prime. On the assumption that the priming effects of Strong and Alternative primes have to do with the salience of relevant alternatives, as the Salience Hypothesis contends, we would then expect alternative priming to have larger priming effects than implicature priming. But this is the opposite of what we observed.

The results of Experiment 2 pose further issues for the Salience Hypothesis. In Experiment 2, we compared two different types of alternative primes for *ad hoc* implicatures, namely, SAME-alternative primes, whose linguistic stimuli were the actual alternatives involved in the generation of the implicatures in the following target trials, and DIFFERENT-alternative primes, which involved the same constructions but noun phrases that were irrelevant for the implicatures in the following target trials. The results show that there is no qualitative difference in the magnitude of their priming effects. In particular, for the X&Y-ALT primes, both types of alternative primes led to quite robust boosting effects. This is unexpected under the Salience Hypothesis. According to this hypothesis, because SAME-alternative prime trials directly mention the crucial alternative to be used later, they should make it more salient than DIFFERENT-alternative prime trials, and consequently the former should have larger boosting effects than the latter. The fact that DIFFERENT-alternative primes were found to robustly increase the rate of implicature responses strongly suggests that the salience of an alternative expression itself is *not* what is driving the priming effect, which would directly contradict the Salience Hypothesis. In other words, what is triggering the priming effect must be more abstract.

We argue that our theory based on context-adaptation provides a more natural explanation of the above findings. Specifically, for cases where the baseline rate is low, strong and alternative priming can lead to a boosting effect, insofar as it provides a cue that a QuD that makes the quantity implicature relevant is more likely to be the intended QuD than otherwise. For Strong primes, it is natural to assume that they serve as strong evidence that the current QuD is one where the quantity implicature is intended, because they force it. In addition, it is reasonable to assume that Alternative primes provide similar cues about the current QuD for all three expressions we tested. Specifically, for SOME, Alternative primes involve *all*, which naturally contrasts with *some*. Having seen utterances involving *all* before seeing *some* could affect one's expectation about what the current QuD is by prompting one to seek for a common QuD that both statements are good answers to. Concretely, two QUDs are conceivable: an open-QUD asking *what proportion of symbols are stars* and a *yes/no*-QUD asking *whether all symbols on the card are stars*. Either way, these QuDs make the 'not all' implicature of the SOME statement in the target highly relevant. Similarly, for NUMBER, the linguistic stimuli of Alternative primes involve *six*, and having seen it before seeing *four* on the target should trigger similar reasoning about the possible common QuDs: an open-QuD asking about *how many stars are on the card* and a *yes/no*-QuD asking *whether there are six stars on the card*. An exact interpretation of the target *four* is naturally expected to answer either QUD.

The same logic applies to *ad hoc* implicatures. In the results of Experiments 1 and 2, we observed reliable priming effects on simple 'X-sentences' of the form *There is a(n) X* with X&Y-ALT prime trials while equally complex X-ALT prime trials had no effect. This is naturally explained by the context-adaptation theory. First, recall that very low implicature-rates were observed in the AD HOC baselines. This suggests that the prior expectation is that the most likely QuD is one that makes *ad hoc* implicature irrelevant, e.g., a *yes/no*-QUD about the existence of X-symbols. Second, since X-ALT primes involve the same construction as AD HOC sentences in the target trials, seeing them will not cause any particular change in one's prior expectations about what kind of QuDs these sentences are likely to be associated with. On the other hand, X&Y-ALT primes involve a more complex X&Y-sentence such as *There is a square and a star*, and this can affect one's reasoning about the likely QuD for the AD HOC sentence in the target trial in a similar way as the other alternative-primes above. That is, having seen X&Y-sentences in the prime phase, one seeks for a common QuD that both these sentences and the AD HOC sentence of the target trial can be addressing. It could be a QuD asking *what is on the card* or a *yes/no*-QuD asking *whether there is a square and a star*. Either way, the *ad hoc* implicature of the X-statement is highly relevant. The results of Experiment 3 shed further light on this question. The baseline rate of *ad hoc* implicatures is higher for X&Y-sentences than for X-sentences. Assuming the Context Adaptation Hypothesis, this can be seen as evidence that in comparison to simple, non-conjunctive sentences, conjunctive sentences have stronger prior associations with QuDs that make *ad hoc* implicatures relevant.

It is furthermore an advantage of our context-adaption approach to priming that it provides a natural explanation for the fact that alternative priming has less robust priming effects than strong priming. This is because alternative priming is typically a less reliable cue than strong priming that the QuD intended for a target trial is one that makes the quantity implicature relevant. More precisely, in the case of strong priming, prime and target trials involve the same scalar expression (if not the same nouns), and according to the way in which expectations are formed in the current context-adaptation model, the likelihood is conditional on the linguistic stimulus. Consequently, a prime trial involving *some*, for example, will have a stronger priming effect on a trial with *some* (to the extent that the prime trial goes against the initial expectation), in comparison to an alternative prime that involves *all* because, in this case, it is relatively more likely that the *all* statement is simply addressing a different QuD. Similar remarks apply to the case of NUMBER. For *ad hoc*, it should be noted that it is not the nouns that matter, but the complexity of the overall sentence, as the results of Experiment 2 and Experiment 3 suggest. This, in turn, means that relevant expectations are about sentence types, rather than about particular lexicalisations of them. This makes sense given that forming expectations about very particular nouns, or more generally about content words, will be generally useless since their frequencies are relatively low and one generally would not expect to encounter the exact same content words in the exact same constructions any time soon, if ever again. On the other hand, expressions like *some*, numeral expressions, and construction types like X-sentences, and X&Y-sentences, are frequent enough and having expectations about how they are likely to be interpreted will facilitate their interpretations. Therefore, the lack of difference between SAME- and DIFFERENT-alternative primes on AD HOC observed in Experiment 2 is consistent with the context-adaptation approach.

9. Concluding remarks

The main focus of our studies has been the nature of priming effects previously reported, since these have been taken as real evidence for a role for salience of alternatives in determining availability of quantity implicatures. To the extent that our results have implications for general language use, the evidence points to a limited role for salience of alternative expressions in deriving implicatures, independently of their impact on expectations about context.

This conclusion has wider implications for linguistic theories of what linguistic expressions can serve as alternatives. One influential theory of alternatives outlined in Katzir (2007) holds that the space of possible alternatives is syntactically delimited by structural modification of what is uttered by syntactic replacements with lexical items as well as linguistic structures that are salient in the discourse (see also Fox & Katzir, 2011). The alternatives so generated are assumed to be further narrowed down to the contextually relevant ones in reference to the QuD and possible other pragmatic factors, before being used to generate quantity implicatures. Importantly, the role of salience in this theory is not merely pragmatic in that it feeds the syntactic generation of alternatives. Our experimental findings and conclusions certainly do not directly undermine this view, but it is still notable that we found no evidence in our studies that the salience of alternatives independently affects the derivation of quantity implicatures. In particular, in the case of the X-alt in Experiment 2 and the X&Y-alt in Experiment 3, we found no additional boost to implicature availability, above and beyond any accounted for by context adaptation. This should give us pause for thought whether salience has any distinctive role in deriving quantity implicatures. To be sure, we have tested these competing ideas within the controlled setting of a laboratory experiment, with multiple trials of the same type, and so it remains to be seen whether similar results emerge from investigations in settings with richer information about context. But that is a question for future research.

Data availability

Stimulus materials, data files and analysis code associated with this research are available open access on OSF at https://osf.io/6gsv9/.

Acknowledgements

This research was supported by the Leverhulme Trust, United Kingdom grant RPG-2018-425.

Appendix A. Instructions for experiments 1 and 2

General — In this study, we will ask for your judgements about English sentences. Every sentence that you will see will be accompanied by two pictures. Your task is to decide which of the two pictures you think the sentence is describing. The study has two parts, Part 1 and Part 2, which slightly differ from one another. Please read carefully the instructions provided to you before you start each part.

Part 1 — Every sentence will be accompanied by two pictures: one of them will be visible to you, while the other one will remain covered with the label ‘Better picture?’ on it. The sentence is meant to describe **one and only one of these two pictures**. Your task is to decide which picture you think the sentence is describing: the visible one or the covered one? You will click on the visible picture if you consider it a match for the sentence; otherwise, you will click on the covered picture.

Part 2 — As in Part 1, every sentence will be accompanied by two pictures. In some cases, one of them will remain covered just as before but, in others, both pictures will be visible to you. As before, the sentence is meant to describe **one and only one of these two pictures** and your task is to decide which picture you think the sentence is describing. You will click on the picture that you consider a better match for the sentence.

References

Antoniou, Kyriakos, Cummins, Chris, & Katsos, Napoleon (2016). Why only some adults reject under-informative utterances. *Journal of Pragmatics*, *99*, 78–95. http://dx.doi.org/10.1016/j.pragma.2016.05.001.

Anwyl-Irvine, Alexander L., Massonnié, Jessica, Flitton, Adam, Kirkham, Natasha, & Evershed, Jo K. (2019). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 388–407. http://dx.doi.org/10.3758/s13428-019-01237-x.

Atlas, Jay David, & Levinson, Stephen (1981). It-clefts, informativeness, and logical form: Radical pragmatics (revised standard version). In Peter Cole (Ed.), *Radical pragmatics* (pp. 1–62). New York: Academic Press.

Bar-Lev, Moshe, & Fox, Danny (2020). Free choice, simplification, and innocent inclusion. *Natural Language Semantics*, *28*, 175–223. http://dx.doi.org/10.1007/s11050-020-09162-y.

Barner, David, Brooks, Neon, & Bale, Alan (2011). Accessing the unsaid: The role of scalar alternatives in children’s pragmatic inference. *Cognition*, *118*(1), 84–93. http://dx.doi.org/10.1016/j.cognition.2010.10.010.

Barr, Dalea J., Levy, Roger, Scheepers, Christoph, & Tily, Harrya J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001.

Bates, Douglas, Mächler, Martin, Bolker, Ben, & Walker, Steve (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

Bates, Douglas, Maechler, Martin, & Bolker, Ben (2011). lme4: Linear mixed-effects models using S4 classes.

Bergen, Leon, Levy, Roger, & Goodman, Noah (2016). Pragmatic reasoning through semantic inference. *Semantics & Pragmatics*, *9*(20), 1–83. http://dx.doi.org/10.3765/sp.9.20.

Bott, Lewis, & Chemla, Emmanuel (2016). Shared and distinct mechanisms in deriving linguistic enrichment. *Journal of Memory and Language*, *91*, 117–140. http://dx.doi.org/10.1016/j.jml.2016.04.004.

Bott, Lewis, & Frisson, Steven (2022). Salient alternatives facilitate implicatures. *PLOS One*, *17*(3), Article e0265781. http://dx.doi.org/10.1371/journal.pone.0265781.

Bott, Lewis, & Noveck, Ira A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language*, *51*(3), 437–457. http://dx.doi.org/10.1016/j.jml.2004.05.006.

Breheny, Richard (2008). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics*, *25*(2), 93–139. http://dx.doi.org/10.1093/jos/ffm016.

Breheny, Richard, Klinedinst, Nathan, Romoli, Jacopo, & Sudo, Yasutada (2018). The symmetry problem: current theories and prospects. *Natural Language Semantics*, *26*(2), 85–110. http://dx.doi.org/10.1007/s11050-017-9141-z.

Büring, Daniel (2003). On D-trees, beans, and B-accents. *Linguistics and Philosophy*, *26*, 511–545. http://dx.doi.org/10.1023/A:1025887707652.

Catani, Marco, & Bambini, Valentina (2014). A model for social communication and language evolution and developpment (SCALED). *Current Opinion in Neurobiology*, *28*, 165–171. http://dx.doi.org/10.1016/j.conb.2014.07.018.

Chemla, Emmanuel (2009). Presuppositions of quantified sentences: experimental data. *Natural Language Semantics*, *17*(4), 299–340. http://dx.doi.org/10.1007/s11050-009-9043-9.

Cremers, Alexandre, Wilcox, Ethan, & Spector, Benjamin (2023). Exhaustivity and anti-exhaustivity in the RSA framework: Testing the effect of prior beliefs. *Cognitive Science*, (47), e13286. http://dx.doi.org/10.1111/cogs.13286.

Fine, Kit (2017). Truthmaker semantics. In Bob Hale, Crispin Wright, & Alexander Miller (Eds.), *A companion to the philosophy of language* (2nd ed.). (pp. 556–577). Hoboken, NJ: Wiley Blackwell.

Fine, Alex B., Jaeger, T. Florian, Farmer, Thomas A., & Ting, Qian (2013). Rapid expectation adaptation during syntactic comprehension. *PLOS One*, *8*(10), Article e77661. http://dx.doi.org/10.1371/journal.pone.0077661.

Foppolo, Francesca, Guasti, Maria Teresa, & Chierchia, Gennaro (2012). Scalar implicatures in child language: Give children a chance. *Language Learning and Development*, *8*(4), 365–394. http://dx.doi.org/10.1080/15475441.2011.626386.

Foppolo, Francesca, Mazzaggio, Greta, Panzeri, Francesca, & Surian, Luca (2021). Scalar and ad-hoc pragmatic inferences in children: guess which one is easier. *Journal of Child Language*, *48*(2), 350–372. http://dx.doi.org/10.1017/S030500092000032X.

Fox, Danny (2007). Free choice and the theory of scalar implicatures. In Uli Sauerland, & Penka Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 71–112). Palgrave Macmillan.

Fox, Danny, & Katzir, Roni (2011). On the characterization of alternatives. *Natural Language Semantics*, *19*(1), 87–107. http://dx.doi.org/10.1007/s11050-010-9065-3.

Frank, Michael C., & Goodman, Noah D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998. http://dx.doi.org/10.1126/science.121863.

Franke, Michael, & Jäger, Gerhard (2016). Probabilistic pragmatics, or why Bayes' rule is probably important for pragmatics. *Zeitschrift für Sprachwissenschaft*, *35*(1), 3–44. http://dx.doi.org/10.1515/zfs-2016-0002.

Geurts, Bart (2006). Take 'five': The meaning and use of a number word. In Svetlana Vogeleer, & Liliane Tasmowski (Eds.), *Non-definiteness and plurality* (pp. 311–329). John Benjamins, http://dx.doi.org/10.1075/la.95.16geu.

Geurts, Bart (2010). *Quantitiy implicatures*. Cambridge: Cambridge University Press.

Geurts, Bart, & Pouscoulous, Nausicaa (2009). Embedded implicatures?!? *Semantics and Pragmatics*, *2*(4), 1–34. http://dx.doi.org/10.3765/sp.2.4.

Grice, Herbert P. (1975). Logic and conversation. In *Speech acts* (pp. 41–58). Brill.

Guasti, Maria Teresa, Chierchia, Gennaro, Crain, Stephen, Foppolo, Francesca, Gualmini, Andrea, & Meroni, Luisa (2005). Why children and adults sometimes (but not always) compute implicatures. *Language and Cognitive Processes*, *20*(5), 667–696. http://dx.doi.org/10.1080/01690960444000250.

Hartigan, John A., & Hartigan, Pamela M. (1985). The dip test of unimodality. *The Annals of Statistics*, 70–84.

Hartsuiker, Robert J., & Kolk, Herman H. J. (1998). Syntactic persistence in dutch. *Language and Speech*, *41*(2), 143–184. http://dx.doi.org/10.1177/002383099804100202.

Hartsuiker, Roibert J., & Westenberg, Casper (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*(2), B27–B39. http://dx.doi.org/10.1016/S0010-0277(99)00080-3.

Horowitz, Alexandra C., Schneider, Rose M., & Frank, Michael C. (2018). The trouble with quantifiers: exploring children's deficits in scalar implicature. *Child Development*, *89*(6), e572–e593. http://dx.doi.org/10.1111/cdev.13014.

Huang, Yi Ting, & Snedeker, Jesse (2009a). Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface. *Cognitive Psychology*, *58*(3), 376–415. http://dx.doi.org/10.1016/j.cogpsych.2008.09.001.

Huang, Yi Ting, & Snedeker, Jesse (2009b). Semantic meaning and pragmatic interpretation in 5-year-olds: evidence from real-time spoken language comprehension. *Developmental Psychology*, *45*(6), 1723.

Hunt, Lamar, Politzer-Ahles, Stephen, Gibson, Linzi, Minai, Utako, & Fiorentino, Robert (2013). Pragmatic inferences modulate N400 during sentence comprehension: Evidence from picture–sentence verification. *Neuroscience Letters*, *534*, 246–251. http://dx.doi.org/10.1016/j.neulet.2012.11.044.

Jaeger, T. Florian (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Jaeger, T. Florian, & Snider, Neal E. (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, *127*, 57–83. http://dx.doi.org/10.1016/j.cognition.2012.10.013.

Kao, Justine T., Wu, Jean Y., Bergen, Leon, & Goodman, Noah D. (2014). Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences*, *111*(33), 12002–12007.

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

Lenth, Russell, Love, Jonathon, & Herve, Maxime (2017). Package 'emmeans'.

Maechler, Martin (2013). Package 'diptest'. R Package Version 0.75–5.

Maldonado, Mora, Chemla, Emmanuel, & Spector, Benjamin (2017). Priming plural ambiguities. *Journal of Memory and Language*, *95*, 89–101. http://dx.doi.org/10.1016/j.jml.2017.02.

Maldonado, Mora, Chemla, Emmanuel, & Spector, Benjamin (2019). Revealing abstract semantic mechanisms through priming: The distributive/ collective contrast. *Cognition*, *182*, 171–176. http://dx.doi.org/10.1016/j.cognition.2018.09.009.

Marty, Paul, Chemla, Emmanuel, & Spector, Benjamin (2013). Interpreting numerals and scalar items under memory load. *Lingua*, *133*, 152–163. http://dx.doi.org/10.1016/j.lingua.2013.03.006.

Marty, Paul, Chemla, Emmanuel, & Spector, Benjamin (2015). Phantom readings: the case of modified numerals. *Language, Cognition and Neuroscience*, *30*(4), 462–477. http://dx.doi.org/10.1080/23273798.2014.931592.

Matsumoto, Yo (1995). The conversational condition on Horn scales. *Linguistics and Philosophy*, 21–60.

Meyer, Marie-Christine, & Feiman, Roman (2021). Priming reveals similarities and differences between three purported cases of implicature: Some, number and free choice disjunctions. *Journal of Memory and Language*, *120*, http://dx.doi.org/10.1016/j.jml.2020.104206.

Myslín, Mark, & Levy, Roger (2016). Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition*, *147*, 29–56. http://dx.doi.org/10.1016/j.cognition.2015.10.021.

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

Noveck, Ira A., Chierchia, Gennaro, Chevaux, Florelle, Guelminger, Raphaelle, & Sylvestre, Emmanuel (2002). Linguistic-pragmatic factors in interpreting disjunctions. *Thinking & Reasoning*, *8*(4), 297–326. http://dx.doi.org/10.1080/13546780244000079.

Noveck, Ira A., & Posada, Andres (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*(2), 203–210. http://dx.doi.org/10.1016/S0093-934X(03)00053-1.

Palan, Stefan, & Schitter, Christian (2018). Prolific.ac — A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.

Pankratz, Elizabeth, & van Tiel, Bob (2021). The role of relevance for scalar diversity: a usage-based approach. *Language and Cognition*, *13*(4), 562–594. http://dx.doi.org/10.1017/langcog.2021.13.

Papafragou, Anna, & Musolino, Julien (2003). Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition*, *86*(3), 253–282. http://dx.doi.org/10.1016/S0010-0277(02)00179-8.

Pouscoulous, Nausicaa, Noveck, Ira A., Politzer, Guy, & Bastide, Anne (2007). A developmental investigation of processing consts in implicature production. *Language Acquisition*, *14*(4), 347–375. http://dx.doi.org/10.1080/10489220701600457.

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

Raffray, Claudine N., & Pickering, Martin J. (2010). How do people construct logical form during language comprehension? *Psychological Science*, *21*(8), 1090–1097. http://dx.doi.org/10.1177/0956797610375446.

Rees, Alice, & Bott, Lewis (2018). The role of alternative salience in the derivation of scalar implicatures. *Cognition*, *176*, 1–14. http://dx.doi.org/10.1016/j.cognition.2018.02.024.

Roberts, Craige (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, *5*, 1–69. http://dx.doi.org/10.3765/sp.5.6.

Romoli, Jacopo (2012). Obligatory scalar implicatures and relevance: a problematic case. *Snippets*, *25*, 11–12.

Scheepers, Christoph (2003). Syntactic priming of relative clause attachments: persistence of structural configuration in sentence production. *Cognition*, *89*(3), 179–205. http://dx.doi.org/10.1016/S0010-0277(03)00119-7.

Schuster, Sebastian, & Degen, Judith (2020). I know what you're probably going to say: Listener adaptation to variable use of uncertainty expressions. *Cognition*, *203*, Article 104285. http://dx.doi.org/10.1016/j.cognition.2020.104285.

Schwarz, Florian (2007). Processing presupposed content. *Journal of Semantics*, *24*(4), 373–416. http://dx.doi.org/10.1093/jos/ffm011.

Schwarz, Florian, & Tiemann, Sonja (2017). Presupposition projection in online processing. *Journal of Semantics*, *34*(1), 61–106. http://dx.doi.org/10.1093/jos/ffw005.

Skordos, Dimitrios, & Papafragou, Anna (2016). Children's derivation of scalar implicatures: Alternatives and relevance. *Cognition*, *153*, 6–18. http://dx.doi.org/10.1016/j.cognition.2016.04.006.

Southgate, Victoria, Chevallier, Coralie, & Csibra, Gergely (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*(6), 907–912. http://dx.doi.org/10.1111/j.1467-7687.2009.00946.x.

Spotorno, Nicola, Koun, Eric, Prado, Jérôme, van der Henst, Jean-Baptiste, & Noveck, Ira A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony. *NeuroImage*, *63*(1), 25–39. http://dx.doi.org/10.1016/j.neuroimage.2012.06.046.

Stiller, Alex, Goodman, Noah, & Frank, Michael (2011). Ad-hoc scalar implicature in adults and children. *vol. 33*, In *Proceedings of the annual meeting of the cognitive science society*. (33).

Stiller, Alex J., Goodman, Noah D., & Frank, Michael C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, *11*(2), 176–190. http://dx.doi.org/10.1080/15475441.2014.927328.

Sun, Chao, & Breheny, Richard (2022). The role of alternatives in the interpretation of scalars and numbers: Insights insights from the inference task. *Semantics and Pragmatics*, *15*, http://dx.doi.org/10.3765/sp.15.8.

Trinh, Tue, & Haida, Andreas (2015). Constraining the derivation of alternatives. *Natural Language Semantics*, *23*(4), 249–270. http://dx.doi.org/10.1007/s11050-015-9115-y.

van Rooij, Robert, & Schulz, Katrin (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information*, *13*(4), 491–519. http://dx.doi.org/10.1007/s10849-004-2118-6.

van Rooij, Robert, & Schulz, Katrin (2006). Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy*, *29*(2), 205–250. http://dx.doi.org/10.1007/s10988-005-3760-4.

Van Rooy, Robert (2003). Questioning to resolve decision problems. *Linguistics and Philosophy*, *26*, 727–763. http://dx.doi.org/10.1023/B:LING.0000004548.98658.8f.

van Tiel, Bob, van Miltenburg, Emiel, Zevakhina, Natalia, & Geurts, Bart (2016). Scalar diversity. *Journal of Semantics*, *33*(1), 137–175. http://dx.doi.org/10.1093/jos/ffu017.

van Tiel, Bob, Pankratz, Elizabeth, Marty, Paul, & Sun, Chao (2019). Scalar inferences and cognitive load. In Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally, & Cristina Real-Puigdollers (Eds.), *Proceedings of Sinn und Bedeutung 23* (pp. 427–441).

van Tiel, Bob, Pankratz, Elizabeth, & Sun, Chao (2019). Scales and scalarity: Processing scalar inferences. *Journal of Memory and Language*, *105*, 93–107. http://dx.doi.org/10.1016/j.jml.2018.12.002.

Waldon, Brandon, & Degen, Judith (2020). Symmetric alternatives and semantic uncertainty modulate scalar inference. In Stephanie Denison, Michael Mack, Yang Xu, & Blair C. Armstrong (Eds.), *Proceedings of the 42nd annual meeting of the cognitive science society*.

Wilson, Deirdre, & Sperber, Dan (1986). On defining relevance. In Richard E. Grandy, & Richard Warner (Eds.), *Philosophical Grounds of Rationality: Intentions, Categories, Ends*, 243–258.