Unless, exceptives, and the pragmatics of conditionals

Prerna Nadathur
Stanford University

Daniel Lassiter
Stanford University

Abstract   The connective *unless* has resisted a compositional semantic analysis, appearing to contribute a biconditional meaning when composed with the positive universal quantifier *every*, but a one-directional meaning when composed with *no*. We report on two experimental investigations comparing the meaning and interpretation of quantified *unless*- and *if not*-conditionals, the results of which argue against the received 'exceptive' treatments of *unless* (von Fintel 1993, 1994, Leslie 2009). The first experiment, using the quantifiers *every* and *no*, demonstrates that *unless* is not semantically biconditional, and suggests that it is instead sensitive to a prohibition against use in contexts where the conditional consequent holds across the board. The second experiment, using the quantifiers *most*, *some*, and *few*, demonstrates contra previous authors that *unless* composes meaningfully with non-universal quantifiers, and clarifies the interpretation of some puzzling results of Experiment 1. Building on these results, we propose a new account of *unless* on which it shares the asserted content of *if not*. Both kinds of conditionals are associated with an inference that the speaker has a reason for avoiding the use of the simpler, unconditional proposition *q*, but this inference is conventionally encoded as a presupposition in the case of *unless*. We argue that the stronger tendency of *unless* toward a biconditional interpretation is due to the interaction of its presuppositional content with a conditional perfection implicature that applies to both kinds of conditionals. We explore the consequences of this difference for the pragmatic behaviour of conditional statements, and suggest some directions for further investigation of quantified statements, conditionals, and exceptive constructions.

1   Introduction

The connective *unless* has often been cited as a potential counterexample to a compositional theory of semantics, on the grounds that it contributes a different meaning when embedded under positive quantifiers than it does when embedded under negative ones (Janssen 1997, Szabó 2008). Although Higginbotham's (1986) original argument relies on the problematic classical identification of *unless* with the negative material conditional *if not* (*q unless p* := $\neg p \rightarrow q$),[1] the intuition that there is a difference between positive and negative quantificational contexts has persisted, and is upheld in more recent work (e.g. Leslie 2009). The most up-to-date form of the problem can be formulated as follows: *unless* appears to contribute a biconditional meaning in positive contexts, but a one-directional conditional in negative ones. An account of *unless* is consequently of some interest with

---

1 This definition appears in Reichenbach (1947) and Quine (1959); see Égré & Cozic (2012) for an overview of arguments against treating natural language conditionals as material implication.

respect to the status of compositionality: the challenge is to develop a semantic account which not only reflects the perceived strength of *unless*-conditionals relative to *if not*-conditionals, but also captures the contextual split in (1).

(1)   a.   Every student will succeed unless he goofs off.
          All students who do not goof off succeed, and all students who goof off do not succeed.
      b.   No student will succeed unless he works hard.
          No student who does not work hard succeeds [but hard work does not guarantee success].

The best-known approach to this problem treats *unless* as an exceptive operator on quantifier domains von Fintel (1992, 1994). In its most current form (due to Leslie 2009), the exceptive account handles the question of compositionality by exploiting formal set-theoretical differences between positive and negative universal quantifiers to build the biconditionality/unidirectionality split directly into the semantics of *unless*. This account makes a number of predictions which have yet to be thoroughly examined, both regarding the behavior of *unless* as compared to *if not*, as well as with respect to the interpretation of *unless* when embedded under non-universal quantifiers.

In this paper, we report first on an experimental test of the predictions for *unless* vs. *if not* in sentences involving the quantifiers *every* and *no*. The results of this study lead to a number of challenges for existing exceptive theories: they suggest not only that semantic biconditionality is too strong a requirement for *unless* in the positive cases, but also that unidirectionality is not strong enough to account for certain contexts in which the use of an *if not* conditional is acceptable, but an *unless* conditional is not.

Based on these results, and on data from a follow-up study of *unless* and *if not* embedded under the non-universal quantifiers *most*, *some*, and *few*, we argue for two main conclusions. First, while the perception of a split between *unless*-biconditionality under positive quantifiers and *unless*-unidirectionality under negative quantifiers reflects an important empirical reality, the difference should not be modeled as a difference in the asserted content of *unless* in these two quantificational contexts. Rather, it should be explained on the basis of pragmatic considerations. Consequently, *unless* does not pose an issue for semantic compositionality. Second, our data suggest that there is a closer relationship between the semantic representations of *if not* and *unless* than proposed by the exceptive accounts. We set out a revised theoretical proposal for *unless*, following the line sketched in Nadathur (2014): *unless* shares the asserted content of *if not*, but differs crucially from *if not* in the relationship it bears to a strengthening inference that is associated with all conditional connectives. Roughly speaking, use of a conditional statement q COND p suggests that the speaker is unable or unwilling to assert the unqualified proposition q: we argue that *unless* lexically encodes a presupposition prohibiting its use in contexts that do not readily justify this inference. While *if not* statements generate a similar inference, it arises in this case as a conversational implicature. The final section of this paper sets out this proposal in detail, and discusses its consequences for future work on conditionals, exceptive operators, and the connection between presupposition, pragmatic inference, and notions of focus or expectation associated with (quantified) generalizations.

2   The exceptive account

2.1   *Unless* and exceptionality

Based on the similarity between (2a) and (2b), *unless* has classically been defined as equivalent to the negative material conditional *if not* (e.g. Quine 1959). Embedded under a negative quantifier, this produces a patently incorrect interpretation (Higginbotham 1986).[2]

(2)   a.   John will succeed unless he goofs off.
       b.   John will succeed if he does not goof off.

In addition to the logical problem, several alternative proposals suggest that this equivalence is inadequate: Clark & Clark (1977) define *q unless p* as *q only if not p*, while Fillenbaum (1986) proposes that it is best represented by *not p only if q*. Crucially, while these proposals contain elements of the negative conditional *if not*, they share the intuition that *unless* is stronger in some way. Dancygier (1985) spells this out more clearly, arguing that, in using *q unless p*, a speaker intends to assert *q*, but also to say *not q if p*. Similarly, Geis (1973) draws attention to a number of data points which seem to suggest that *unless* patterns more closely with *except if* than with *if not*.

The central idea that emerges is that *unless*-statements not only condition *q* on *not p* but simultaneously draw attention to the truth of *q* in those cases where *p* does hold. Von Fintel (1992) formalizes this by treating *unless* as an exceptive operator on quantifier domains (akin to *except for* in "Every student except for John attended the meeting"). On this account, an *unless*-statement conjoins two assertions. First, it generalizes over a domain from which the *unless*-complement is subtracted. Second, secondly, that this complement represents the unique smallest set over which the generalization fails.

(3)   **Analysis based on von Fintel's 1992 proposal:**[3]
       $Q[C]M \text{ unless } R := Q[C-R]M \wedge \forall S \subseteq C : Q[C-S]M \rightarrow R \subseteq S$

(4)   Every student will succeed unless he goofs off.
       $\text{ALL}[\text{STUDENT} - \text{GOOF}]\text{SUCC} \wedge \forall S \subseteq \text{STUDENT} : \text{ALL}[\text{STUDENT} - S]\text{SUCC} \rightarrow \text{GOOF} \subseteq S$

This results in the interpretation in (4) for (1a), where STUDENT represents the set of students, GOOF the set of individuals who goof off, and SUCC the set of those who succeed. The first

---

2 This is a familiar problem for material conditionals, as illustrated by (i), which reduces to the claim that everyone neither succeeds nor works hard:

   (i)   No one will succeed unless he works hard.
          $= \neg \exists x (\neg \text{work-hard}(x) \rightarrow \text{succeed}(x))$
          $= \neg \exists x (\text{work-hard}(x) \vee \text{succeed}(x))$
          $= \forall x \neg (\text{work-hard}(x) \vee \text{succeed}(x))$
          $= \forall x (\neg \text{work-hard}(x) \wedge \neg \text{succeed}(x))$

3 Von Fintel's original proposal is stated for cases where Q is a modal quantifier. He does not provide an explicit formula interpreting *unless*-statements with a nominal quantifier, so there is some uncertainty as to how to adapt the proposal for these cases. Leslie (2009) shows that allowing the quantifier to scope over a (covert) universal modal quantifier results in the same problem as the classical account. Given these considerations, (3) seems to us to be the most plausible extension of von Fintel's proposal to the cases at hand; moreover, it parallels a proposal for *if* in von Fintel (1998).

conjunct asserts that all students who do not goof off are successful. The second clause (von Fintel's 'uniqueness') produces the effect of the reverse conditional (all Ms are not Rs), by stipulating that students who goof off are necessarily excluded from any arbitrary set which contains only successful students. This entails that no students who goof off succeed, and a von Fintel-style exceptive account thus produces a biconditional interpretation for (4).

2.2 Capturing the negative cases

Leslie (2009) points out that proposal (3) also predicts biconditionality in negative quantificational contexts:

(5) No student will succeed unless he works hard.
$$\text{NO}[\text{STUDENT} - \text{WORK}]\text{SUCC} \ \wedge \ \forall S \subseteq \text{STUDENT}: \ \text{NO}[\text{STUDENT} - S]\text{SUCC} \rightarrow \text{WORK} \subseteq S$$

That is, no student who does not work hard will succeed, and no student who does work hard is contained in any set of unsuccessful students. Consequently, working hard is not only a necessary condition for success, but also guarantees it.

Leslie argues that biconditionality is too strong here, and there is good reason to take this objection seriously. Suppose, for instance, we are discussing a university course that is notoriously difficult. We know that students taking this class must work very hard to pass, but it may be the case that for some of them, even hard work is not enough to guarantee success. In a situation where none of the students who do not work hard succeed, but even some of those who work very hard do not pass, (5) seems to be neither invalid nor infelicitous. This suggests that the semantics of (5) should only require that working hard is a necessary condition for success, and crucially not that it is sufficient. Leslie proposes to capture this distinction (non-sufficiency under the negative quantifier) by modifying the uniqueness clause as shown (bold) in (6):

(6) **Leslie's proposal:**[4]
$$Q[C]M \ unless \ R := Q[C-R]M \wedge \mathbf{Q}[\mathbf{C} \cap \mathbf{M}](-\mathbf{R})$$

Since the negative universal quantifier *no* is symmetric – that is, *No As are Bs* is logically equivalent to *No Bs are As* – this modification has the desired result of preserving biconditionality under *every* while eliminating it under *no*, as shown in (7).

(7)   a. Every student will succeed unless he goofs off.
$= \text{ALL}[\text{STUDENT} - \text{GOOF}]\text{SUCC} \wedge \text{ALL}[\text{STUDENT} \cap \text{SUCC}](-\text{GOOF})$
All students who don't goof off succeed, and all students who succeed don't goof off.
   b. No student will succeed unless he works hard.
$= \text{NO}[\text{STUDENT} - \text{WORK}]\text{SUCC} \wedge \text{NO}[\text{STUDENT} \cap \text{SUCC}](-\text{WORK})$
No student who doesn't work hard succeeds, and no student who succeeds doesn't work hard.
Equivalently: *No student who doesn't work hard succeeds.*

---

4 Some of the details of Leslie's account have been glossed over here. Crucially, Leslie holds that *unless* can restrict nominal quantifiers as well as quantificational adverbs, and this is reflected in (7). The motivation for this is discussed in §6.1.

2.3 Issues with the exceptive account

Despite the elegance of Leslie's solution, naturally-occurring examples of *unless* conditionals reveal several issues with her version of the exceptive account. These issues can be divided into two classes. §2.3.1 deals with the interaction of uniqueness with a special kind of context (referred to as 'across-the-board' contexts), and illustrates the problem with non-universal quantifiers. §2.3.2 demonstrates the issues that arise from stipulating uniqueness as an entailed consequence of *unless*.

2.3.1 Non-universal quantifiers and across-the-board contexts

Von Fintel (1992) claims that *unless* can only co-occur with universal quantifiers. His semantics provide a natural explanation for this restriction: the notion of a unique exception from, e.g., a *most*-generalization is incoherent. Leslie (2009) disagrees with this restriction on empirical grounds, and her view is supported by the existence of a large number of naturally-occurring examples of *unless* under non-universal quantifiers, including those in (8).

(8)   a. "Most Americans won't go to church unless they have a need." [One News Now 2015]
      b. "California is so dry, some diners won't get water unless they ask." [NPR 2014]
      c. "Few people can be happy unless they hate some other person, nation, or creed." [attr. to Bertrand Russell in Prochnow (1955)]
      d. "Smoking kills half of smokers unless they quit." [Gates Foundation 2014]

Statements of this sort ultimately pose a problem for Leslie's account as well as von Fintel's. Consider the interpretation that Leslie's theory assigns to (9).

(9)   Most students will pass unless they skip class.
      $= \text{MOST}[\text{STUDENT} - \text{SKIP}] \text{ PASS} \land \text{MOST}[\text{STUDENT} \cap \text{PASS}](-\text{SKIP})$
      Most students who do not skip class pass and most students who pass do not skip class.

At first glance, this seems reasonable, but things go wrong if the domain of students under consideration is dominated by students who skip class. Scenario A in Figure 1 illustrates. If 9 of these students pass, satisfying the first conjunct, the second conjunct then requires that these 9 comprise *most* (at least more than half) of the students who pass over the whole domain of 110. Thus, at most 8 (and possibly fewer) of the 100 students who skip class can pass without falsifying (9). This seems much too strong a condition to impose. In particular, (9) is an intuitively reasonable description of a parallel situation in which 10, or 15 (or any number up to at least half) of the 100 students who skip class still pass; that is, of any situation in which perhaps some but crucially *not most* of the students who skip class pass. Leslie's formula rules out any situations where the number of class-skippers who pass is greater than or equal to the number of class-attendees who pass.

Scenario A suggests that the truth conditions of Leslie's second conjunct are too strong. However, there are also contexts in which the truth conditions of (9) seem to be too general: consider Scenario B in Figure 1. We have 12 students, only 4 of whom skip class. 6 of the students who do not skip class pass, and 3 others also pass Then we have 6 students who pass among the 8 who do not skip class, which satisfies the first conjunct of (9). We also have 6 students who do not skip class among

5

the total of 9 who pass, so the second conjunct is also satisfied. In this situation, however, in addition to the fact that most of the students who attend class pass, it also turns out that most of the students who skip class pass. Thus, skipping class makes absolutely no difference to the student pass rate: 75% of students pass, whether they skip class or not. Here, (9) seems to be a highly inappropriate description of the situation, but Leslie's formula makes no provision for this.
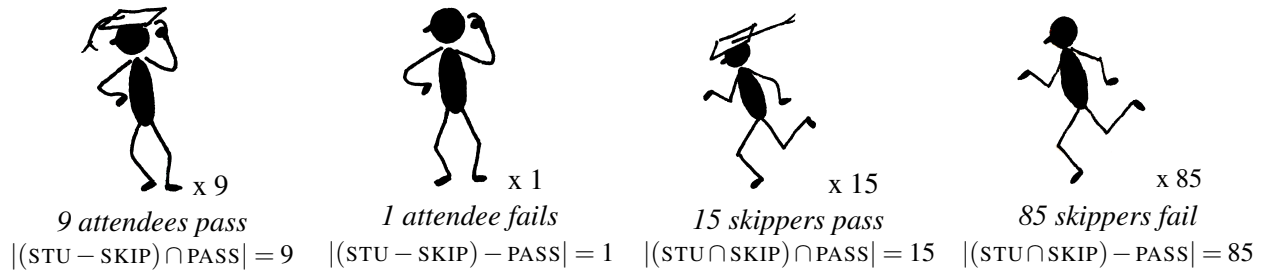
Scenario A

Most students will pass unless they skip class is intuitively **true**

110 students total $|\text{STUDENT}| = 110$

10 students do not skip   *100 students skip*

$|\text{STUDENT} - \text{SKIP}| = 10$ $|\text{STUDENT} \cap \text{SKIP}| = 100$

| x 9 | x 1 | x 15 | x 85 |

9 attendees pass   *1 attendee fails*   *15 skippers pass*   *85 skippers fail*

$|(\text{STU} - \text{SKIP}) \cap \text{PASS}| = 9$ $|(\text{STU} - \text{SKIP}) - \text{PASS}| = 1$ $|(\text{STU} \cap \text{SKIP}) \cap \text{PASS}| = 15$ $|(\text{STU} \cap \text{SKIP}) - \text{PASS}| = 85$

The second conjunct of Example (9) (Leslie 2009) is **false**

$$\text{MOST}[\text{STUDENT} \cap \text{PASS}](\neg \text{SKIP}) = F$$

because $|(\text{STUDENT} \cap \text{PASS}) - \text{SKIP}| < |(\text{STUDENT} \cap \text{PASS}) \cap \text{SKIP}|$
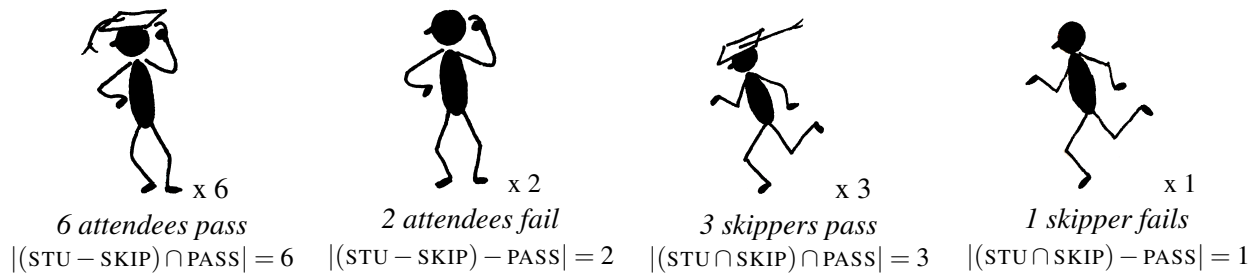
Scenario B

Most students will pass unless they skip class is intuitively **false**

12 students total $|\text{STUDENT}| = 12$

8 students do not skip   *4 students skip*

$|\text{STUDENT} - \text{SKIP}| = 8$ $|\text{STUDENT} \cap \text{SKIP}| = 4$

| x 6 | x 2 | x 3 | x 1 |

6 attendees pass   *2 attendees fail*   *3 skippers pass*   *1 skipper fails*

$|(\text{STU} - \text{SKIP}) \cap \text{PASS}| = 6$ $|(\text{STU} - \text{SKIP}) - \text{PASS}| = 2$ $|(\text{STU} \cap \text{SKIP}) \cap \text{PASS}| = 3$ $|(\text{STU} \cap \text{SKIP}) - \text{PASS}| = 1$

Both conjuncts of Example (9) (Leslie 2009) are **true**

$\text{MOST}[\text{STUDENT} - \text{SKIP}]\text{PASS} = \top$ $\text{MOST}[\text{STUDENT} \cap \text{PASS}](\neg \text{SKIP}) = \top$

because $|(\text{STUDENT} - \text{SKIP}) \cap \text{PASS}| >$ because $|(\text{STUDENT} \cap \text{PASS}) - \text{SKIP}| >$

$|\text{STUDENT} - \text{SKIP}) - \text{PASS}|$ $|(\text{STUDENT} \cap \text{PASS}) \cap \text{SKIP}|$

Figure 1: Two problematic contexts for Leslie (2009).

Truth conditions that seem too general are not inherently problematic. We might appeal to pragmatics to rule out the problematic scenarios above. This does not help with a related problem associated with (6), however: the second conjunct of Leslie's proposal yields what seem to be overly specific truth conditions in cases with non-universal quantifiers. Consider (10), which gives Leslie's truth conditions for (8d):

(10) HALF [SMOKER − QUIT] DIE ∧ HALF [SMOKER ∩ DIE] (−QUIT)   (=8d)
      Half of the smokers who do not quit die and half of the smokers who die do not quit.

Suppose we have a total of $2N$ smokers who never quit. Interpreting HALF loosely as *approximately half*, (10) requires that roughly N of these non-quitting smokers die (from smoking-related causes). This number must also be approximately equal to half the number of former smokers who die from smoking. Thus, we have $\sim N$ smokers who do not quit and are fine, $\sim N$ smokers who don't quit and die as a consequence, and $\sim N$ smokers who quit but die anyway. Intuitively, however, (8d) should not specify anything about the relative size of the intersecting domains of smokers who don't quit and one-time smokers who die from smoking. It should be possible to truthfully use (8d) regardless of the relative size of these sets, but Leslie's truth conditions require them to be roughly equivalent.

The problem is exacerbated if we couple this conclusion with a common-sense inference ruling out the type of scenario that caused the generality problem for *most*. Specifically, if we assume that (8d) would not be used in a situation where quitting smoking has no statistical effect on death rate, we are forced to conclude that the number L of quitters who reap the rewards of quitting and do not die from smoking-related causes is either (significantly) less than or greater than N. World-knowledge tells us that (8d) is a warning, suggesting that it would not be used in a situation where quitting increases the rate of smoking-related death. This gives us $L >> N$. The upshot of Leslie's truth conditions is a prediction that (8d) should license the inference that significantly more people quit smoking than not (since $L + N >> 2N$). This conclusion does not seem at all warranted.

The problem illustrated by (9)-(10) is that, while the reformulated uniqueness condition gives seemingly reasonable predictions for quantifiers like *every* and *no*, it makes increasingly odd predictions for quantifiers that are non-universal, by restricting the type and construction of the domains under discussion in nonintuitive ways. Moreover, as Scenario B for (9) illustrates, we have a clear intuition that *unless* statements should be (at least) infelicitous in **across-the-board** contexts: situations where the main generalization ($Q[C]M$) holds on the entire domain, including the set (R) marked as an exception. For universally-quantified statements, the across-the-board cases are ruled out by uniqueness, since they logically contradict biconditionality. However, the uniqueness clause fails to capture the relevant condition under non-universal quantifiers. Insofar as the naturally-occurring examples in (8) support the view that *unless* is both grammatical and interpretable with non-universal quantifiers, this represents a serious flaw in the exceptive account.

2.3.2 Biconditionality as pragmatic inference

Having observed that a uniqueness clause causes problems in across-the-board contexts as well as negative quantificational ones, it seems sensible to ask whether this clause ought to have been

included in the semantics of *unless* in the first place. Pushing further on the idea that biconditionality is not always the correct interpretation for *unless*, we find naturally-occurring examples that suggest that biconditionality is not entailed even by positive and universally quantified examples:

(11) Mantou is always late unless she's already out before we meet, but she's often just less late then.

On both exceptive proposals, the main clause in (11) requires that all relevant situations are (a) such that Mantou is late if she is not already out, and (b) that she is not late otherwise. However, the *but*-clause specifies that at least some of the situations where Mantou is out are ones in which she is still late, albeit less so. Thus, on the exceptive theories, (11) ought to appear as contradictory as (12). However, it seems completely acceptable.

(12) #Roses are always red and violets are always blue, but sometimes violets are not blue.

This difference shows that uniqueness/biconditionality is defeasible, and thus that it cannot be entailed by *unless*. Crucially, however, if we suppress the *but*-clause in (11), we perceive a strong tendency to interpret the example biconditionally. This suggests that uniqueness is pragmatically associated with *unless*, at least in positive quantificational contexts.

Nadathur (2014) provides additional evidence for treating uniqueness as an implicature, which we reiterate briefly here. First, (13a) shows that uniqueness (the second conjunct in both von Fintel and Leslie's formulas) can be reinforced without apparent redundancy: entailments, by comparison, do not have this property.

(13) a. "Always be yourself, unless you are Fernando Torres. Then always be someone else."
      b. *Compare:* Always be yourself, unless you are Fernando Torres. ?Otherwise always be yourself.

Second, example (14a) shows that uniqueness can be questioned without contradiction: this contrasts with the results of questioning entailed content.

(14) a. "The answer is no unless you ask. If you do ask the answer might still be no."
      b. *Compare:* The answer is no unless you ask. #If you don't ask the answer might be yes.

The uniqueness inference therefore displays defeasibility (see 11), reinforceability, and question-ability, which are properties of pragmatic inferences. Nadathur (2014) also provides (15)-(16). The former demonstrates that uniqueness is backgroundable, arguing against a conventional implicature treatment (see Potts 2005). (Note that the (a) examples are more natural with prosodic focus on *unless*, presumably due to independent pressure to mark the discourse contrast between *if* and *unless*.) The latter shows that explicitly suspending uniqueness does not cause infelicity, arguing against a presuppositional account.

(15) a. John won't fail if he studies. He will fail unless he studies.
      b. *Compare:* John is a student. John, ?the student, will fail unless he studies.

(16) a. The student might not fail if he studies, but he will fail unless he studies.
      b. *Compare:* ?There might not be a student, but he will fail unless he studies.

Together with this evidence, the pervasiveness of the tendency to interpret *unless* statements biconditionally speaks to a characterization of the uniqueness inference as a 'default' pragmatic inference, in keeping with Levinson's (2000) characterization of generalized conversational implicatures (GCIs). It corresponds to our intuitions about 'normal' interpretations for *unless*-conditionals, and moreover bears a striking resemblance to the well-known GCI of *conditional perfection* (the tendency for *if*-conditionals to be interpreted biconditionally; Geis & Zwicky 1971), which also displays defeasibility, questionability, and reinforceability. (17) is Geis and Zwicky's original example of conditional perfection.

(17) I'll give you five dollars if you mow the lawn.
      ⤳ *I'll give you five dollars if and only if you mow the lawn.*

Both inferences (conditional perfection and uniqueness) are *nonconventional*, in the sense that they are not lexically encoded and do not attach in all circumstances – consider, for instance, the one-directional negative contexts for *unless* and (18) for conditional perfection.

(18) If this cactus grows native to Idaho, it's not an *Astrophytum*.
      ⤴ *If and only if this cactus grows native to Idaho, it's not an* Astrophytum

The precise characterization of conditional perfection is a matter for independent research (for an overview of past research, see van der Auwera 1997), but the parallel between this and uniqueness emphasizes the default nature of biconditional interpretation in both cases; we thus tentatively classify uniqueness, alongside perfection, as a GCI.

3   Experiment 1: *unless* and universal quantifiers

We are faced with a puzzling collection of intuitions about *unless*. There is evidently an empirical difference between *unless* and *if not*, which seems to reside in the fact that *unless*-statements draw attention to the truth of their main generalization over the excepted set (R). It is unclear, however, what precisely is communicated in this regard. In the preceding section, we argued that a biconditionality/uniqueness inference is cancelable, and is therefore not an entailment of *unless*. On the other hand, if it is merely an implicature, it is not immediately clear why it should associate with *unless* more strongly than with *if not*. Moreover, the discussion in §2.3.1 suggests that the exceptive account is missing something, insofar as *unless*-statements are apparently infelicitous in across-the-board contexts (those in which the main generalization holds on both the excepted set and its complement). Making the puzzle more complex, we also need to account for the divergence between *unless* under *every* and under *no*. Having rejected a semantic uniqueness clause in favour of a pragmatic account, we can no longer fall back on Leslie's explanation of the asymmetry.

This section reports on an experiment with two main motivations.[5] First, as §2 demonstrates, any analysis based only on a small set of intuitions is empirically limited. For one, such intuitions do not distinguish easily between semantic and pragmatic aspects of interpretation. Secondly, as the examples in (19) show (with $Q=every$), intuitions about appropriate and inappropriate uses of

5 These results were previously reported in Nadathur & Lassiter (2014).

9

unless are robust only in 'edge' cases, where either *Q*-many of the individuals in the excepted set do not have the property picked out by the nuclear scope (biconditional), or *Q*-many of them do (across-the-board).

(19) [Context: *Half of the students goofed off.*] Every student passed unless he goofed off.

    a. **Clearly appropriate** if all of the students who did not goof off passed, and all of the students who goofed off did not pass. (Biconditional context; uniqueness satisfied)

    b. **Clearly inappropriate** if all of the students passed, including all of those who goofed off. (Across-the-Board context; uniqueness not satisfied)

    c. **Unclear/??** if all of the students who did not goof off passed, and {a few/half/most/...} of the students who goofed off passed. (Intermediate context; uniqueness not satisfied)

The intermediate cases are equally unclear when *Q=no*. Collecting native speaker judgements on a large scale under controlled conditions seems to be the only way to clarify these cases. Further, as discussed below, the quantitative details of these cases are highly informative about the status of exceptionality with *unless* and the nature of the difference between *unless* and *if not*. These details ultimately place significant constraints on the parameters of a revised theoretical account of *unless*.

3.1   Design

Participants were shown a display of 20 red and blue marbles, and were asked to decide whether a given stimulus statement about the display was true or false (forced-choice; Figure (3.1) is a sample trial). Test stimuli (examples 20-21) contained either *unless* or *if not*, embedded under either *every* or *no*. We varied the proportion of target-colour marbles with dots from among 0, 0.2, 0.4, 0.6, 0.8, and 1. This gave us a total of 24 test conditions.
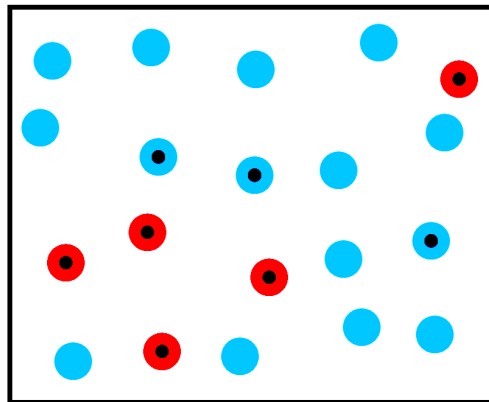
Is the following claim true or false?

"Every marble has a dot unless it is blue."

◯ True   ◯ False

Continue

Figure 2: Sample test trial

(20) a. Every marble has a dot unless it is [target colour].

   b. Every marble has a dot if it is not [target colour].

(21) a. No marble has a dot unless it is [target colour].

   b. No marble has a dot if it is not [target colour].

To increase display variety, we randomly varied the target colour between red and blue, and the ratio of target:non-target marbles between 5:15, 10:10, and 15:5 (each test condition thus had 6 display variants). To avoid overwhelming participants with false sentences, we set the proportion of dotted non-target marbles in each test condition to satisfy the minimal truth conditions of both *unless* and *if not*, as given by $Q[C - R]M$ (the first conjunct in (2) and (5)).

 We also included a number of fillers. A 'sampling' condition asked participants to imagine that a marble was chosen at random, and then judge a sentence of the form "The selected marble has a dot {if it is not/unless it is} [colour]." The display varied as described above.[6] Additional filler statements varied along three parameters: quantifier (*every, no*, no quantifier), whether they mentioned 'red,' 'blue,' or did not mention colour at all, and construction type (22-24). For the last parameter, we used positive *if*-sentences, single-clause quantified statements, and *there*-existentials. In these filler displays, we also varied the red:blue ratio from 5:15, 10:10, and 15:5, and selected both red and blue dot proportions randomly from 0, 0.2, 0.4, 0.6, 0.8, and 1. Consequently, any given filler statement could occur with any one of 108 possible displays.

(22) [no quantifier, red, *if*]: The selected marble has a dot if it is red.

(23) [*every*, dot, single-clause]: Every marble has a dot.

(24) [*no*, blue, *there*]: There are dots on no blue marbles.

3.2 Method

Using Amazon's Mechanical Turk platform, we recruited 160 participants, all of whom were financially compensated. They viewed the experiment as a HIT (Human Intelligence Task) through the Mechanical Turk website. Participants were given detailed instructions (accompanied by a sample display and a non-test stimulus), and were told that in each trial they would be presented with a different display-stimulus pair and asked to judge whether the stimulus was true or false of the display. They were unable to proceed from one trial to the next without selecting an answer. At the end of the experiment, participants were asked to report their native language; this did not affect payment for the HIT.

 Each participant saw 48 randomly-ordered trials. 24 of these trials were randomly selected from a set containing the 24 test conditions and 12 'sampling' conditions. The displays for these were selected, again at random, from the 6 variants for the given condition. 24 trials were randomly selected from the remaining 27 filler types, with one of the 108 display variants randomly generated for each filler. Each test item was seen by an average of 105 participants (min=80, max=124, median=109).

---

6 We included this filler condition out of curiosity about 'sampling'-based judgments. The results were statistically indistinguishable from those in the *every* test condition, consistent with theories in which bare (unquantified) indicative conditionals contain a covert *must* (Kratzer 1986). We do not analyze these results further here.

3.3 Results

We excluded data from 5 participants who reported being native speakers of languages other than English. The analysis below includes data from the remaining 155 participants.
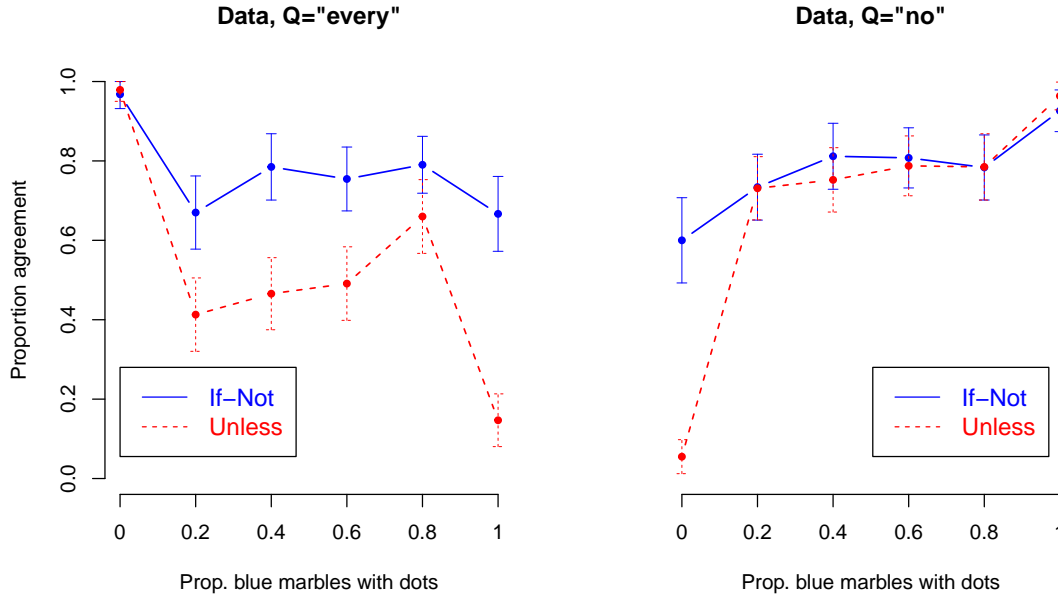

Figure 3: Main results of Experiment 1. Error bars represent 95% confidence intervals.

Figure 3 shows our results in both quantifier conditions (left=*every*; right=*no*). In each graph, the *x*-axis represents the proportion of target marbles with dots, and the *y*-axis represents the fraction of participants who judged the sentence true. Results from the *if not* condition are in blue with a solid line interpolated between points. Results from the *unless* condition are in red with a dashed line. Error bars represent 95% binomial confidence intervals. Table 1 gives the corresponding numerical data.

| | | if not | | | | | unless | | |
| | | *every* | | *no* | | | *every* | | *no* |
Proportion	N	% agree	N	% agree	N	% agree	N	% agree
0.0	93	96.8 ± 3.6	80	60.0 ± 10.7	95	97.9 ± 2.9	109	5.5 ± 4.3
0.2	100	67.0 ± 9.2	109	73.4 ± 8.3	109	41.3 ± 9.2	119	73.1 ± 8.0
0.4	93	78.5 ± 8.4	85	81.2 ± 8.3	116	46.6 ± 9.1	109	75.2 ± 8.1
0.6	110	75.5 ± 8.0	104	80.8 ± 7.6	112	49.1 ± 9.3	113	78.8 ± 7.5
0.8	124	79.0 ± 7.2	97	78.4 ± 8.2	100	66.0 ± 9.3	93	78.5 ± 8.4
1.0	96	66.7 ± 9.4	95	92.6 ± 5.3	109	14.7 ± 6.6	110	96.4 ± 3.5

Table 1: Endorsement rates in test conditions.

We analyzed the data using a separate linear mixed-effects models for each quantifier, using the lme4 package (Bates et al. 2014) in R (R Core Team 2014). Proportion was coded as a categorical variable, and we included random effects of participant, target colour, and red/blue distribution (5:15, 10:10, 15:5). We included random intercepts only because the maximal models with random slopes did not converge. For each quantifier, we tested for a main effect of conditional type (*if not* vs. *unless*) following the procedure outlined by Levy (2014). Specifically, we converted the categorical proportion variable to a sum-coded numeric variable, and then calculated the likelihood ratio of two models differing only in the inclusion of a fixed main effect of conditional type. Both models included a fixed main effect of proportion and an interaction between proportion and context, in addition to random intercepts as noted above. These tests revealed a highly significant main effect of conditional type for both quantifiers. For *every*: $\chi^2(1) = 33.3$, $p < 10^{-8}$. For *no*: $\chi^2(1) = 12.4$, $p < .001$.

To provide a baseline rate for erroneous responses, Table 2 summarizes responses to filler stimuli paired with displays which clearly falsified them. Participants very rarely responded 'true' in such scenarios (37 of 1549 judgments in total; 2.4%).

QUANT *marble is [colour]*		existential variant		QUANT *marble has a dot*		existential variant	
N	% agree	N	% agree	N	% agree	N	% agree
538	1.8 ± 1.2	456	3.5 ± 1.9	276	2.2 ± 1.9	269	1.8 ± 1.8

Table 2: Endorsement rates for false filler items

3.4 The qualitative patterns and their implications for previous theories

This section discusses the predictions of von Fintel's and Leslie's theories and compares them to our experimental results. We find that the results are not consistent with the predictions of either previous theory.

Both formulations (3, von Fintel; 6, Leslie) of the exceptive account make unambiguous predictions about the truth-values of the relevant *unless* sentences (25-26). Table 3 shows the predicted distribution of truth-values by theory in each experimental condition.

(25) Every marble has a dot unless it is blue.
von Fintel/Leslie: TRUE iff *all red marbles have dots and no blue marbles have dots.*

(26) No marble has a dot unless it is blue.
von Fintel: TRUE iff *no red marbles have dots and all blue marbles have dots.*
Leslie: TRUE iff *no red marbles have dots.*

| | von Fintel | | | | | | Leslie | | | | | |
| | Target dot proportion | | | | | | Target dot proportion | | | | | |
	0.0	0.2	0.4	0.6	0.8	1.0	0.0	0.2	0.4	0.6	0.8	1.0
every	T	F	F	F	F	F	T	F	F	F	F	F
no	F	F	F	F	F	T	T	T	T	T	T	T

Table 3: Predictions for *unless* by condition and theory

13

In general, we assume that sentences where *if not* restricts a nominal quantifier have the interpretation (27). The corresponding predictions are given in (28), and are assumed to hold across all theories considered here.

(27) $Q[C]M$ *if not* $R := Q[C - R]M$

(28)   Every/no marble has a dot if it is not blue.
       TRUE iff *all/no red marbles have dots.*

Since we designed the experiment so that the truth-conditions for *if not* sentences were satisfied in all test conditions, we predict a response of TRUE in all cases involving *if not* (see Table 4). Results from this condition should thus provide a baseline for interpreting *unless* results.

| | **if not** | | | | | |
| | Target dot proportion | | | | | |
	0.0	*0.2*	*0.4*	*0.6*	*0.8*	*1.0*
every	T	T	T	T	T	T
no	T	T	T	T	T	T

Table 4: Predictions for *if not* by condition

We discuss results for *no* first. Under *no*, participants almost unanimously endorsed stimuli with both *if not* and *unless* at target proportion 1. Recall that the non-target dot proportion in the *no* condition was set at 0: target proportion 1 therefore represents the biconditional context in which no non-target marbles have dots and all target marbles do (satisfying uniqueness). This result is consistent with both von Fintel's and Leslie's accounts.

Endorsement rates dropped off slightly at proportions less than 1 but greater than 0, and appear to have done so identically for *unless* and *if not*. If we allow that the lowered acceptability is due to some pragmatic pressure on both connectives, this result is consistent with Leslie's account, on which *unless* and *if not* are semantically identical under *no* (see Tables 3 and 4). However, it is not consistent with the von Fintel-style theory (3), which predicts that the *unless* sentences are always biconditional, and thus false under *no* when the proportion of target marbles with dots is less than 1 (Table 3). Since our participants categorically rejected false items (only 2.4% acceptance, see Table 2), the high rates of endorsement in this condition falsify theory (3).

The remaining data point under *no* is problematic for Leslie's account: *unless* diverges from *if not* sharply at target proportion 0. While a majority of participants (60%) found *No marble has a dot if it is not blue* acceptable in the 'Across-the-Board' scenario where no marbles of either colour had dots, endorsement of the corresponding *unless*-sentence approached zero (5.5%) in the same scenario. Indeed, the divergence at 0 was so sharp that it appears to have been entirely responsible for the significant main effect of connective under *no* reported above: endorsement rates for the two connectives were indistinguishable at all other proportions. Confirming this impression, a model comparison procedure identical to the one outlined above, but with proportion 0 removed from the data in all conditions, found no main effect of connective choice under *no* ($\chi^2(1) = .01, p > .9$). Since Leslie predicts identical truth-values for *unless* and *if not* across all target proportions under *no* (see Tables 3-4), this divergence is inconsistent with her theory.

Results under *every* were in some respects the mirror image of those under *no*, but there are some important differences. Here again we found near-unanimous endorsement in the biconditional context (target proportion 0), just as with *no*. That is, virtually all participants endorsed *Every marble has a dot COND it is blue* for both conditionals when uniqueness/biconditionality was satisfied: all red marbles have dots, and no blue ones do. This result is consistent with both previous semantic accounts. In addition, as under *no*, *if not* and *unless* diverged sharply in the 'Across-the-Board' context (target proportion 1; all marbles have dots regardless of colour). A majority of participants (66.7%) accepted *Every marble has a dot if it is not blue* when all of the pictured marbles had dots, but few (14.7%) accepted *Every marble has a dot unless it is blue* in the same scenario. These data points are consistent with both exceptive accounts, since both predict that *unless* under *every* is true only in biconditional scenarios.

Results in the intermediate range of target proportions (0.2-0.8), however, were problematic for both previous theories. As in the *no* condition, endorsement rates under *every* were non-maximal at these proportions, but remained much higher than endorsement rates for false filler items. However, with *every* there was also a reliable difference between *unless* and *if not* in the intermediate range: *Every marble has a dot if it is not blue* was more likely to be accepted than *Every marble has a dot unless it is blue* when all red marbles have dots and the proportion of blue marbles with dots was .2, .4, .6, or .8. Despite this, it appears that neither statement is false or otherwise totally unacceptable: endorsement of *every-unless* was reliably above floor in this range (between 41% and 66%), while false filler items were very rarely endorsed (3% on average, see Table 2). While von Fintel and Leslie are both correct in arguing that there is a difference between *every-if not* and *every-unless*, the specific truth-conditional diagnosis of this difference appears incorrect: *every-unless* is not simply false in all non-biconditional contexts. In other words, the difference between *unless* and *if not* under *every* appears to involve graded factors affecting felicity, rather than categorical factors involving truth.

In sum, our results argue against a semantically biconditional account of *unless* in the context of either quantifier (*every* and *no*). Such an account would incorrectly predict an overwhelming preference for 'false' responses in non-biconditional contexts (0-0.8 under *no*, 0.2-1 under *every*). The results also argue against Leslie's one-directional account of *unless* under *no*, given the divergence between *unless* and *if not* in Across-the-Board contexts. A one-directional theory cannot explain our participants' near-unanimous rejection of *unless* in these contexts. Finally, the results suggest that there is a non-categorical difference between the two conditionals in intermediate scenarios under *every*. Both previous theories wrongly predict a categorical, truth-conditional difference in these scenarios.

4   Three puzzles and two proposals

The experimental investigation reported in §3 suggests that a new account of *unless* is needed. This account should ideally explain each of the following three puzzles:

(A) The categorical divergence of *unless* from *if not* in Across-the-Board contexts (at target dot proportion 1 under *every*, and at 0 under *no*).

(B) The degraded but non-zero acceptability of both *unless* and *if not* in the middle range of target dot proportions – and, for *if not*, in Across-the-Board contexts as well.

(C) The reliable but non-categorical divergence between *if not* and *unless* under the middle range of target dot proportions in the *every* condition, and the fact that no such divergence appears under *no*.

In this section, we propose solutions for Puzzles (A) and (B), and outline some strategies for explicating Puzzle (C). This discussion is motivated solely by the data reported in the preceding section. In subsequent sections, we will expand on these proposals, both by considering data from a follow-up experiment against the predictions made here, as well as by unifying our suggestions with the theoretical account of *unless* offered in Nadathur (2014).

4.1 Puzzle (A): Categorical divergence in Across-the-Board scenarios

When the quantifier Q is either *every* or *no*, Across-the-Board (AtB) scenarios for statements of the form $Q[C]M$ COND R are easily characterized: they are those in which the simpler statement $Q[C]M$ holds. More specifically, they are contexts in which the sentence obtained by omitting the consequent clause (i.e. COND R) of the conditional produces a true description. For instance, an AtB scenario for "Every marble has a dot unless it is blue" is one in which every marble pictured – including the blue ones – has a dot.

Were it not for Puzzle (A), it would be tempting to suppose that *if not* and *unless* have the same semantic content, and that any differences between them are located strictly in the pragmatics. While there is at least one difference that could in principle be implicated in pragmatic reasoning – *if not* comprises two syntactically separable items, while *unless* is a single lexical unit – we do not know of any well-motivated pragmatic mechanisms that would apply only in AtB contexts and could be expected to cause the observed categorical difference in truth-value judgements. Instead, our results suggest that *unless* conventionally encodes a prohibition against being used in AtB contexts, and that this prohibition is, crucially, absent in the meaning of *if not*.

We refer to this requirement as the **AtB prohibition**. The boldfaced conjuncts in (29)-(30) represent two ways of formulating the AtB prohibition, bracketing for the moment any question of its status as entailment, presupposition, etc.

(29) $Q[C]M$ *unless* $R \Rightarrow Q[C-R]M \land \neg\mathbf{Q[C]M}$ **(Option 1)**

(30) $Q[C]M$ *unless* $R \Rightarrow Q[C-R]M \land \neg\mathbf{Q[C\cap R]M}$ **(Option 2)**

Options 1 and 2 both accurately capture the results comprising Puzzle (A). In fact, (29) and (30) are logically equivalent when Q is either *every* or *no* (assuming that both sets $C-R$ and $C\cap R$ are nonempty). Recalling the naturally-occurring data first reported in §2 (see 31), which suggest that *unless* is compatible with non-universal quantifiers, we believe that Option 2 is to be preferred. As shown in (31), using Option 1 to interpret an *unless*-statement when $Q = $ SOME results in a contradiction. Option 2 on the other hand yields a satisfiable proposition: assuming again the

16

red-blue marble paradigm from Experiment 1, (31b) is true just in case some of the pictured red marbles have dots but none of the blue ones do.[7]

(31)   Some marbles have a dot unless they are blue.

    a. $\text{SOME}[\text{MARBLE} - \text{BLUE}]\text{DOT} \wedge \neg\textbf{SOME}[\textbf{MARBLE}]\textbf{DOT}$ **(Option 1)**

    b. $\text{SOME}[\text{MARBLE} - \text{BLUE}]\text{DOT} \wedge \neg\textbf{SOME}[\textbf{MARBLE} \cap \textbf{BLUE}]\textbf{DOT}$ **(Option 2)**

Representing *unless* – and the AtB prohibition, in particular – as in (30) is, additionally, consistent with results from a follow-up experiment focusing on the quantifiers *most, some*, and *few*. This study is discussed in the next section.

It is important to note that, in the cases at hand, the AtB prohibition makes weaker predictions than the uniqueness clauses formulated by von Fintel and Leslie (3 and 6, respectively). When Q is *every* or *no*, the only effect of the prohibition is to rule out AtB scenarios, as desired. Biconditionality/uniqueness, on the other hand, is falsified also in contexts where intermediate proportions of the excepted set satisfy the scope of the quantifier.[8]

(32)   Every marble has a dot unless it is blue.

    a. *If not +* **AtB prohibition**:
       $\text{ALL}[\text{MARBLE} - \text{BLUE}]\text{DOT} \wedge \neg\textbf{ALL}[\textbf{MARBLE} \cap \textbf{BLUE}]\textbf{DOT}$
       Every red marble has a dot, and some blue marble does not.

    b. *If not +* **uniqueness**:
       $\text{ALL}[\text{MARBLE} - \text{BLUE}]\text{DOT} \wedge \textbf{ALL}[\textbf{MARBLE} \cap \textbf{BLUE}]\neg\textbf{DOT}$
       Every red marble has a dot, and every blue marble does not.

(33)   No marble has a dot unless it is blue.

    a. *If not +* **AtB prohibition**:
       $\text{NO}[\text{MARBLE} - \text{BLUE}]\text{DOT} \wedge \neg\textbf{NO}[\textbf{MARBLE} \cap \textbf{BLUE}]\textbf{DOT}$
       No red marble has a dot, and some blue marble does.

    b. *If not +* **uniqueness**:
       $\text{NO}[\text{MARBLE} - \text{BLUE}]\text{DOT} \wedge \textbf{NO}[\textbf{MARBLE} \cap \textbf{BLUE}]\neg\textbf{DOT}$
       No red marble has a dot, and no blue marble does not.

Given the results of Experiment 1, it seems clear that some version of the AtB prohibition is lexically associated with *unless*, replacing uniqueness. The remaining question is thus whether the AtB prohibition represents an entailment, presupposition, conventional implicature, or something

---

7 The discussion in §8.5 of Peters & Westerståhl (2008) addresses a similar problem, in this case for exceptive expressions such as *except* and *but for* (e.g., "Every student except the chemists attended the meeting"). Option 1 for the AtB prohibition closely corresponds to von Fintel (1993)'s proposal for the meaning of 'free' (unconnected) exceptives, while Option 2 resembles more closely an older proposal from Hoeksema (1987, 1990) (examples (NC2) and (NC1), respectively, in Peters & Westerståhl; p.302). Our support for Option 2, on the basis of its compatibility with non-universal quantifiers, thus amounts to the view that, if *unless* is to be seen as an exceptive, it must be in the sense of Hoeksema, rather than the stricter class postulated by von Fintel.

8 For exceptive constructions, again, Peters & Westerståhl (2008) describe the difference between an anti-AtB formulation ($\neg Q[C \wedge R]M$) and a von Fintel-style uniqueness clause ($Q[C \wedge R]\neg M$) as the difference between 'outer' and 'inner' negation. This relation ultimately appears to be relevant to Puzzles (B) and (C).

else altogether. Empirical endorsement rates for *unless* statements in AtB scenarios are comparable to the rates at which truth-conditionally false filler statements were endorsed; however, since our experimental design did not include control conditions with false presuppositions or failed conventional implicatures, these numbers represent, at best, circumstantial evidence that the AtB prohibition should be modeled as an entailment of *unless*. Moreover, results from Abrusán & Szendröi (2013), Schwarz (2016), and, most recently, Cremers et al. (2016) suggest that sentences carrying failed presuppositions are often treated by naïve subjects in the same manner as truth-conditionally false statements. Consequently, the low acceptance rates for *unless* statements in AtB contexts are consistent with an interpretation of the AtB prohibition as a presupposition or precondition associated with felicitous use of *unless*.

Independent arguments suggests that the AtB prohibition is not a conventional implicature (Potts 2005). Example (34) shows that, in contrast with CIs in general, the AtB prohibition can be felicitously backgrounded.

(34)   Some of the blue marbles do not have dots. But every marble has a dot unless it is blue.
       Compare: Some of the blue marbles do not have dots. #The blue marbles, some of which have no dot, are my favourites.

Similarly, reinforcing the AtB prohibition does not result in redundancy. This argues against treating it as an entailment of *unless*.

(35)   Every marble has a dot unless it is blue, and some blue marbles do not have dots.
       Compare: #Every marble has a dot unless it is blue, and every non-blue marble has a dot.

In addition to the arguments from (34) and (35), efforts such as (36a) to explicitly suspend the AtB prohibition result in infelicity. The only exception to this is where the suspension is framed as a correction, as in (36b). This is reminiscent of the behaviour of presuppositional content: compare (36) with (37).

(36)   a.   ??Every blue marble has a dot, and every marble has a dot unless it is blue.
       b.   Every marble has a dot unless it is blue. In fact, every blue marble has a dot as well.

(37)   a.   ??It's not raining, and Mary doesn't realize it's raining.
       b.   Mary doesn't realize it's raining. In fact, it isn't raining at all.

While we remain at this point somewhat uncertain about this diagnosis – due, in part, to the difficulty of applying the standard 'family of sentences' projection tests (Chierchia & McConnell-Ginet 1990) to quantified *unless* statements (and conditional statements as a class), we would like to suggest that the data in (34)-(37), together with our experimental results, favour a presuppositional account of the AtB prohibition. *Modulo* this requirement, we propose that *unless* is semantically/assertorically equivalent to *if not*.

(38) $Q[C]M$ *unless* R $\begin{cases} \text{is a presupposition failure if } Q[C \cap R]M; \text{ otherwise,} \\ \text{is true if and only if } Q[C - R]M. \end{cases}$

Regardless of the precise status of the AtB prohibition, however, it is evident that a condition which sharply distinguishes *unless* from *if not* in AtB contexts must be part of any reasonable account of *unless*. As formulated, the AtB prohibition provides a solution to Puzzle (A), by uniquely picking out those scenarios in which our *unless* test sentences triggered near-zero endorsement rates, in contrast to their *if not* counterparts.

4.2   Puzzle (B): Embedded biconditionality inferences

Puzzle (B) involves the less-than-unanimous but clearly non-zero acceptance rates for both *unless* and *if not* across the middle range of target dot proportions, and in AtB scenarios as well for *if not*. We suggest that this pattern can be explained pragmatically in terms of *conditional perfection* (Geis & Zwicky 1971). As noted in §2.3.2, the biconditional interpretation is usually treated as a generalized conversational implicature (GCI; van der Auwera 1997, Horn 2000, Levinson 2000), which somehow adds the inferred material in (39b) to the asserted content of a conditional like (39a). This results in the interpretation (39c):

(39)   a. The marble Bill selected has a dot if it is not blue.

      b. ⤳ The marble Bill selected does not have a dot if it is blue.

      c. (39a) & (39b) ≡ The marble Bill selected has a dot iff it is not blue.

Conditional statements involving the quantifiers *every* and *no* also seem susceptible to biconditional interpretation, as spelled out in (40)-(41).[9]

(40)   a. Every marble has a dot if it is not blue.

      b. ⤳ Every marble does not have a dot if it is blue.

      c. (40a) & (40b) ≡ All and only non-blue marbles have dots.

---

9 There is an extensive literature on conditional perfection, and a wide range of different approaches to explaining the process by which biconditionality inferences arise. One aspect of this debate centers around whether or not quantified conditionals comprise a structure in which the quantifier directly embeds a conditional operator, or is itself modified by the conditional. The latter corresponds to the Kratzer (1986)-style restrictor analysis adopted by Leslie (2009) (and which we ultimately endorse), while the former is supported by von Fintel & Iatridou (2002), Higginbotham (2003), Huitink (2009), and a new analysis of quantified indicative conditionals from Kratzer (in press). On an embedded conditional analysis, the perfection inference attached to a bare conditional such as (39a) is supposed to arise at the embedded level, and propagate upwards (as an implicature) to give rise to a biconditional which remains embedded under the nominal quantifier. The structure of the inferences in (40b) and (41b) are harder to account for on a restrictor analysis of conditionals: indeed, several authors, including von Fintel (2001) and Franke (2009), discuss the possibility that inferences like (40b) and (41b) arise via a more immediate 'conditional strengthening' (von Fintel) or 'weak conditional perfection' (Franke) inference which is identical to Peters & Westerståhl's 'outer negation.' That is, von Fintel and Franke suggest that the appearence of embedded conditional perfection arises from an inference with the form we have proposed for the AtB prohibition, but which is argued to be generated as an implicature on *if/if not* conditionals. This approach seems particularly promising given the data from Experiment 2. For current purposes, however, it is enough to observe that our empirical evidence points to the reality of inferences of the (b) (embedded *not if*) types, which appear to associate – cancellably – with the conditionals in (40a) and (40b). Whether they arise directly or indirectly, then, we remain concerned here simply with their effect on participant response – and, as we argue, with the fact that the (b) inferences, which give rise to the biconditional (c) interpretations, predict the Puzzle (B) results we observe.

(41) a. No marble has a dot if it is not blue.

      b. ⤳ No marble does not have a dot if it is blue.

      c. (41a) & (41b) ≡ All and only blue marbles have dots.

Previous work has revealed a certain tendency for participants in truth-value judgement tasks to reject true sentences associated with false implicatures. In Doran et al. (2012)'s investigation of true sentences with false GCIs, participants in a baseline condition rejected such sentences at rates ranging between 15% and 63%, over a variety of GCI triggers. In terms of our study, the effect of a default conditional perfection inference, formulated as in (40)-(41), would then be to render *if not* sentences systematically less acceptable in the intermediate target dot proportion, as well as in AtB scenarios, all of which falsify biconditionality under *every* and *no*.

Concretely, suppose that our participants were inclined, with some (small) probability p, to reject *if not* sentences associated with empirically false conditional perfection implicatures. *Modulo* experimental noise, we would then expect an endorsement rate of $(1 - p)$ in the *every-if not* condition for those stimuli where (40a) holds but (40b) does not. Since our experimental design guaranteed that (40a) was satisfied (i.e. all red/non-target marbles had dots) in any test trial in this condition, we expect high but strictly non-maximal endorsement of *every-if not* sentences in any context where the target dot proportion was greater than 0 (that is, for any associated test display where some of the blue/target marbles had dots). For the *if not* test sentences under *no*, we expect endorsement at a rate of $(1 - p)$ when (41a) is true and (41b) is false: again, since all *no-if not* test diplays verified (41a) (no red/non-target marbles had dots), sub-maximal endorsement should be seen just in case the target dot proportion is less than 1 (when at least one pictured blue/target marble does not have a dot). These predictions perfectly describe the results comprising Puzzle (B) as far as *if not* is concerned: our results demonstrate a small but non-trivial tendency for participants to reject *if not* statements in contexts with non-biconditional displays.

Extending this explanation to Puzzle (B) as it pertains to *unless* is straightforward. In (38), we proposed that the asserted content of *unless* is identical to that of *if not* whenever the AtB prohibition is satisfied. It is therefore reasonable to expect that *unless* will also be associated with a conditional perfection inference; §2.2 gives arguments for treating uniqueness/biconditionality effects as pragmatic. In non-AtB contexts, associating a perfection implicature with *unless* will derive the Puzzle (B) patterns. Under *no*, acceptance rates for *unless* are more or less matched with those for *if not* in all non-AtB contexts. We have already dicussed the divergence between the two conditionals in AtB contexts.

An important question remains. While *every-unless* is, as expected, endorsed below ceiling at target dot proportions 0.2-0.8, the explication given for *every-if not* predicts that this should be comparable both to the acceptance rate $(1 - p)$ for *every-if not* over the same proportions, as well as with acceptance for *no-unless* over 0.2-0.8. While the stronger effect observed for *unless* is consistent with the account we have given for Puzzle (B), it leaves something more to be said. This is the subject of Puzzle (C).

4.3 Puzzle (C): Grades of biconditionality

(C) is the most perplexing of our three puzzles. It involves the reliable but non-categorical divergence between *every-if not* and *every-unless* over the target dot proportions 0.2-0.8, as well as the fact that this divergence is realized only under *every*. As noted, endorsement rates for *no-unless* and *no-if not* are statistically indistinguishable in non-AtB contexts. We have grouped these two issues together into a single puzzle because we suspect that their explanations are related. In light of the two proposals we have made thus far, we see two potential approaches to the questions raised by Puzzle (C).

First, higher rejection rates for *unless* under *every* in non-AtB contexts might be attributable to a new pragmatic pressure, distinct from the conditional perfection inference discussed above. This pressure would need to be triggered only in sentences with *every*, and would (presumably) interact additively with the downward pressure exerted by the failure of the perfection implicature in the relevant middle-range contexts. This approach essentially takes as its premise that our experimental results under *no* represent the expected or 'base' case for the behaviour of *unless*, once the AtB prohibition and a conditional perfection/biconditionality implicature are taken into account. From this perspective, what needs to be explained is the special case represented by *every-unless*: specifically, what additional inference or implicature causes for *unless* in the middle range of target dot proportions to be more strongly dispreferred. This possibility is an intriguing one, but on the basis of experimental data alone, we do not have any speculation as to what the nature or contribution of such a pressure might be.

A second possibility is that we are in fact dealing with only one inference – the implicature to biconditionality. It would then be the case that the biconditional implicature is somehow stronger for *unless* than it is for *if not*, but that this strengthening is only present with certain quantifiers. This line of reasoning could privilege either *every* or *no* as the base case: it might be that a regular perfecting inference is made stronger in the presence of *every*, or, alternatively, that an inherently strong implicature towards biconditionality associated with *unless* is made weaker in the presence of *no*. An account of this sort recalls the intuition that we started with (see Leslie 2009), that *unless* appears to be more biconditional in positive than in negative contexts. While our results show that Leslie's specific proposal is incorrect, a pragmatic approach that allows for different strengths of biconditionality would honour the original insight to some extent.

Based only on the experimental data reported in §3, it is difficult to take a strong position regarding the source of Puzzle (C). Ultimately, based on additional data from our follow-up study, as well as theoretical considerations, we believe that the second possibility is more promising (and also provides the impetus for unifying Puzzle (C) as a single problem). Regardless of the ultimate resolution of this question, however, the data and proposals offered so far demonstrate two important things. First, Puzzle (C) is at the heart of the original problem with *unless*: the appearence of a meaning difference between *unless* composed with *every* and *unless* composed with *no*. Second, we have shown that this problem cannot be handled truth-conditionally in the manner previously suggested: *every-unless* cannot entail a biconditional like (40c), as our participants did not reject such statements outright in non-biconditional and non-AtB contexts. Consequently, we must pursue a pragmatic solution to the biconditionality question, and to Puzzle (C). We believe that this conclusion, along with the claim that *if not* and *unless* differ semantically only with respect

to presupposition (i.e. the AtB prohibition), represents significant progress in understanding the conditions governing the meaning and use of *unless*.

5 Experiment 2: *unless* and non-universal quantifiers

So far, we have put forward a view of *unless* as sensitive to the truth and/or relevance of three conditions: an *if not* clause, a prohibition against Across-the-Board contexts, and a *not if* clause:

(42) *Q[C]M unless R*:
    a. *if not* $Q[C \wedge \neg R]M$
    b. *AtB prohibition* $\neg Q[C \wedge R]M$
    c. *perfection* $Q[C \wedge R]\neg M$

(42) clearly entails (42a), and we have argued that it also carries (42b) as a presupposition and frequently conveys (42c) as an additional implicature. Before proceeding with a formalization of the proposal that (42) outlines, however, a few questions bear closer examination.

First, we argued for the AtB presupposition in (42b) over a simpler form based on compatibility with non-universal quantifiers (see 30). *Unless* does felicitously co-occur with these quantifiers (*pace* von Fintel 1993; cf. also Moltmann's 1995 *quantifier condition* for exceptives), as shown in the naturalistic examples from §2 repeated here.

(8) a. "Most Americans won't go to church unless they have a need."
    b. "California is so dry, some diners won't get water unless they ask."
    c. "Few people can be happy unless they hate some other person, nation, or creed."
    d. "Smoking kills half of smokers unless they quit."

In this section, we report on results from an experimental study of *unless* statements with the quantifiers *most*, *some*, and *few*. This study uses the forced-choice paradigm employed previously to test the predictions of our earlier analysis in these more complex cases, and to gain insight into the problematic Puzzle (C) involving the *every/no* split in the middle range of proportions.

5.1 Design

Experiment 2 used the same design as Experiment 1, with the exception of a change in the quantifiers studied and some minor adjustments to the displays and filler and control items.

Participants were shown trials consisting of a display of 20 red and blue marbles, and were asked to judge whether an associate stimulus sentence was true or false of the given display. Test stimuli were determined by four independent variables. *Unless-* and *if not*-sentences were compared in the context of the quantifiers *most, some*, and *few*. The test sentences had the forms in (43)-(45).

(43) a. Most marbles have a dot unless they are [target colour].
    b. Most marbles have a dot if they are not [target colour].
(44) a. Some marbles have a dot unless they are [target colour].

  b. Some marbles have a dot if they are not [target colour].

(45) a. Few marbles have a dot unless they are [target colour].

  b. Few marbles have a dot if they are not [target colour].

The proportion of target-colour marbles with dots varied evenly over $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$. In this study, we also varied the proportion of *non-target* marbles that had dots: each stimulus sentence was assigned to one of the low, mid, or high conditions for non-target dots. The parameters for these conditions were set specifically for the quantifier which appeared in the test sentence: $\{0.6, 0.8, 1\}$ for $\{$low, mid, high$\}$, respectively, in sentences with *most*, $\{0.2, 0.4, 0.6\}$ for *some*, and $\{0, 0.2, 0.4\}$ for *few*. This additional manipulation was included in order to manage uncertainty about the quantifiers' conditions of applicability: we wanted to be able to factor out cases where participants rejected sentences because they did not judge that the proportion of non-target marbles with dots counted as *most*, *few*, or *some*. The baseline acceptability of the quantifiers as a function of proportion was measured in a separate non-conditional control condition, and the non-target proportion with the highest baseline acceptability was selected for analysis in the conditional constructions (see below).

  Taken together, our four variables yielded 2 (conditionals) $\times$ 3 (quantifiers) $\times$ 6 (target dot proportions) $\times$ 3 (non-target dot proportions) = 108 test items. In this experiment, we limited display variation to the random choice of red or blue as the target colour: in all items, the display contained exactly 10 red and 10 blue marbles.

  Filler items varied by quantifier condition (*most, some* or *few*), whether they were *if*-conditionals, single-clause quantified statements, or *there*-existentials. When filler items specified a target colour, it was randomly selected. This results in a total of 3 (quantifier) $\times$ 3 (construction types) = 9 colour-isomorphic filler sentences, each of which could appear with a randomly-generated pair of target and non-target dot proportions for $6 \times 6 = 36$ display variations. Examples of fillers are given in (46):

(46) a. [most, *if*, blue] Most marbles have a dot if they are blue.

  b. [some, single-clause] Some marbles have a dot.

  c. [few, existential] There are dots on few marbles.

  Control items were similar in construction to fillers, but explicitly mentioned either red or blue as a target colour in addition to varying by quantifier and construction type. We restricted display generation for these items to vary evenly between the appropriate low, mid, and high target dot proportions. Non-target dot proportions were selected at random. In total, this yielded 3 (quantifier) $\times$ 3 (construction types) $\times$ 3 (target dot proportions) = 27 control items, each of which could appear with any one of six colour-isomorphic display variants. Examples of our controls are given in (47).

(47) a. [most, *if*, low, red: 60% of red marbles have dots]
   Most marbles have a dot if they are red.

  b. [some, single-clause, mid, blue: 40% of blue marbles have dots]
   Some blue marbles have a dot.

  c. [few, existential, high, red: 40% of red marbles have dots]
   There are dots on few red marbles.

5.2   Method

We recruited 380 participants via Mechanical Turk, all of whom were financially compensated. Participants viewed the experiment as a HIT through the MTurk website. Instructions and debrief questions were the same as for the previous study.

Each participant saw 48 trials in a randomized order. 24 of these were randomly-selected test items, and the remaining 24 were drawn at random from the fillers and controls. Test items were seen by an average of 82.9 participants (min=61, max=103), and control items were seen by an average of 221.6 participants (min=199, max=241).

5.3   Predictions

If we set aside the postulated pragmatic effects of conditional perfection, and assume that each of the low, mid, and high conditions are reasonable representations of the relevant quantifier, we can use the working proposal in (38) to generate predictions for the *unless* test stimuli. Table 5 details the relevant predictions, with low, mid, and high conditions collapsed for each quantifier.

(38) $Q[C]M$ *unless* R $\begin{cases} \text{is a presupposition failure if } Q[C \cap R]M; \text{ otherwise,} \\ \text{is true if and only if } Q[C-R]M. \end{cases}$

	unless					
	Target dot proportion					
	0	0.2	0.4	0.6	0.8	1
most	T	T	T	#	#	#
some	T	#	#	#	#	#
few	#	#	#	T	T	T

Table 5: Predictions for *unless* by condition

Since *if not* simply omits the presuppositional content of (38), and our experimental design satisfies $Q[C-R]M$ in all test trials, regardless of which conditional was used, we expect a positive response to *if not* test sentences in all cases, *modulo* pragmatic effects (e.g., conditional perfection).

5.4   Results

We excluded data from 7 participants who reported being native speakers of languages other than English (including native bilinguals). This left data from 373 participants for analysis. We first present the results from our control stimuli, in order to establish which of the available proportional representations (*low, mid,* or *high*, specified by quantifier), were overall considered to best represent each quantifier. Figure 4 shows the control results by construction type and target dot condition. In each graph (left=*most*, center=*some*, right=*few*), the *x*-axis represents target dot proportion, and the *y*-axis endorsement rates. Results from *if* conditionals (e.g., 47a) are in green with a dashed line

interpolated between points, results from the single-clause quantifier controls (labeled as 'plain') are in purple with solid interpolation, and results from the existential controls are in orange with dotted interpolation. Error bars represent 95% binomial confidence intervals. Table 6 gives the numerical data corresponding to Figure 4, with *N* representing the total number of participants who responded to a given item.
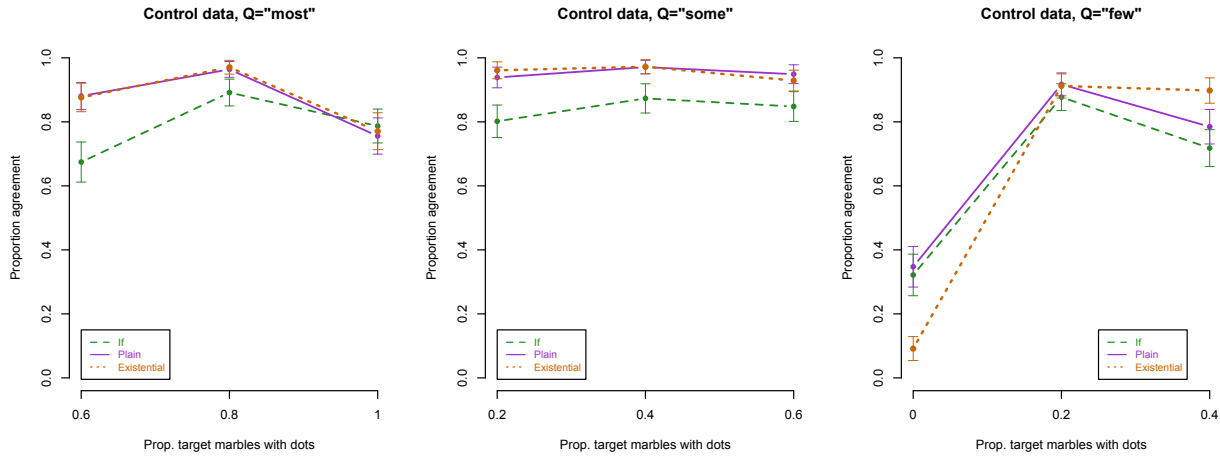

Figure 4: Control results, by quantifier condition

			most		*some*		*few*	
			N	% agree	N	% agree	N	% agree
	if	low	215	67.4 ± 6.3	237	80.2 ± 5.1	199	32.2 ± 6.5
		mid	212	89.2 ± 4.2	205	87.3 ± 4.6	236	87.8 ± 4.2
		high	230	78.7 ± 5.3	230	84.8 ± 4.6	234	72.8 ± 5.8
Target marble condition	plain	low	226	88.1 ± 4.2	212	93.9 ± 3.2	216	34.7 ± 6.4
		mid	221	96.4 ± 2.5	241	97.1 ± 2.1	229	91.8 ± 3.6
		high	221	75.6 ± 5.7	215	94.8 ± 2.9	223	78.5 ± 5.4
	exist.	low	210	87.6 ± 4.5	205	96.1 ± 2.7	229	9.2 ± 3.7
		mid	237	97.1 ± 2.2	216	97.2 ± 2.2	216	91.2 ± 3.8
		high	205	77.1 ± 5.8	239	92.9 ± 3.3	225	89.8 ± 4.0

Table 6: Numerical data for controls

Control data provide information as to the fraction of respondents who accept a given proportion as an adequate representation of the relevant quantifier. In each case, we find that the *mid* condition (target proportion 0.8 for *most*, 0.4 for *some*, and 0.2 for *few*) was endorsed at the highest rate ($\geq 86\%$ agree), regardless of the type of control sentence. This suggests that these conditions are reliably treated as representative of the quantifier in question. Although endorsement for the *some* controls is roughly consistent across all three conditions, the sharper distinction evident in the *most*

and *few* data show that, for these quantifiers, at least, there is disagreement between participants as to whether or not the *low* and *high* target conditions meet the requirements of the relevant quantifier.

Our predictions in the previous section were based on the assumption that the non-target dot condition (whether *high, mid*, or *low*) is a good representation of the quantifier. Consequently, any disagreement or uncertainty with respect to this could represent a potential confound in the interpretation of our results. On the basis of our control results, then, we will confine the analysis in the main text to test conditions in which the entailment conditions of both *if not* and *unless* statements (i.e. 42a) were judged, by the overwhelming majority of participants, to be satisfied. The full data set, including all non-target conditions for each quantifier, is described in the appendix. These results are also in line with the analysis in the main text, though the results in the *high* and *low* conditions are in some cases complicated by the fact that the non-target conditions were not always judged to be good representatives of the quantifier: for example, endorsement of *few*-statements was very low across the board when the non-target proportion was 0.

Figure 5 shows the test results from the *mid* (in this case, corresponding to non-target dot proportions) for each quantifier (left=*most*, center=*some*, right=*few*). In each case, the *x*-axis represents the proportion of target marbles with dots, and the *y*-axis the rate of participant agreement. Results from *unless* are in red with a dashed line interpolated between points, and results from *if not* are in blue with solid interpolation. Error bars represent 95% binomial confidence intervals. Table 7 gives the numerical data corresponding to Figure 5, with *N* as the total number of participants who responded to a particular item.
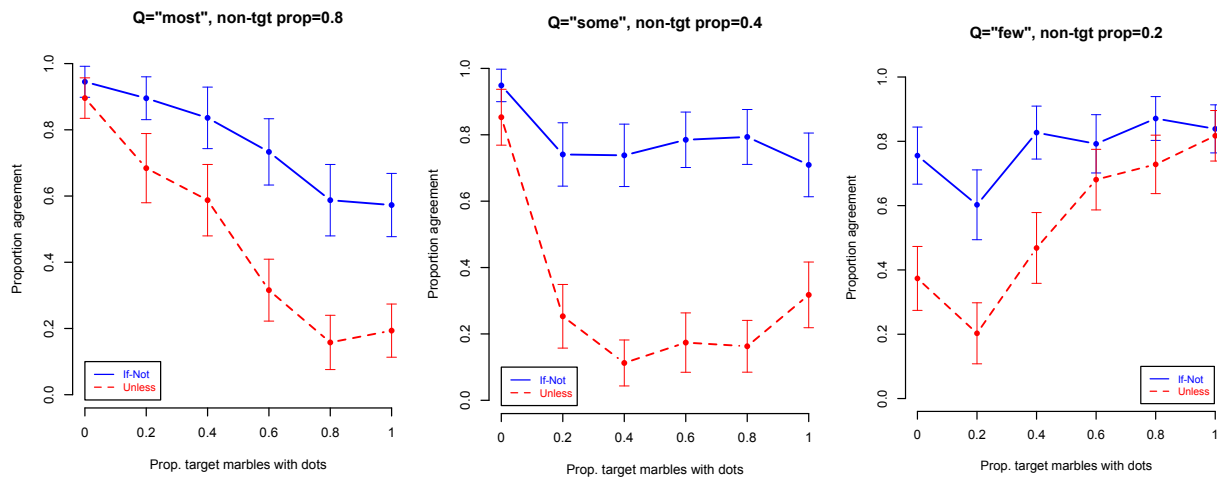

Figure 5: Results from the *mid* condition, by quantifier

			MOST		SOME		FEW	
			N	% agree	N	% agree	N	% agree
		0.0	91	95.5 ± 4.7	78	94.9 ± 4.9	90	75.6 ± 8.9
		0.2	86	89.5 ± 6.5	81	74.1 ± 9.5	78	60.3 ± 10.9
	if not	*0.4*	61	83.6 ± 9.3	84	73.8 ± 9.4	81	82.7 ± 8.2
		0.6	75	73.3 ± 10.0	93	78.5 ± 8.3	77	79.2 ± 9.1
		0.8	80	58.8 ± 10.8	92	79.4 ± 8.3	93	87.1 ± 6.8
Prop.		*1.0*	103	57.3 ± 9.6	86	70.9 ± 9.6	93	83.9 ± 7.5
target								
marbles		*0.0*	96	89.6 ± 6.1	68	85.3 ± 8.4	91	37.4 ± 9.9
with		*0.2*	76	68.4 ± 10.5	79	25.3 ± 9.6	69	20.3 ± 9.5
dots		*0.4*	80	58.8 ± 10.8	80	11.3 ± 6.9	79	46.8 ± 11.0
	unless	*0.6*	95	31.6 ± 9.4	69	17.4 ± 8.9	94	68.1 ± 9.4
		0.8	76	15.8 ± 8.2	86	16.3 ± 7.8	92	72.8 ± 9.1
		1.0	93	19.4 ± 8.0	85	31.8 ± 9.9	93	81.7 ± 7.9

Table 7: Numerical data for *mid* conditions, by quantifier

There is a clear main effect of connective choice in the *mid* conditions in each case. Using the same analytic procedure as Experiment 1 (cf. Levy 2014), there is a highly significant main effect of connective choice for *few* at the optimal non-target proportion .2 ($\chi^2(4) = 86.5, p < 10^{-15}$). The same holds for *some* at non-target proportion .4 ($\chi^2(4) = 322.1, p < 10^{-15}$) and for *most* at non-target proportion .8 ($\chi^2(4) = 145, p < 10^{-15}$).

5.5 Discussion: The Across-the-Board condition

Experiment 2 confirms our claim that *unless* felicitously combines with the non-universal quantifiers *most, some*, and *few*. If these quantifiers were simply incompatible with *unless*, we would expect a high degree of uncertainty in responses to *unless* statements, or perhaps a trend toward unconditional rejection. Instead, the responses are consistent with the theoretical analysis proposed in response to Experiment 1.

The results in Figure 5 and Table 7 are broadly consistent with the predictions of our proposals in §4 and §5.3. First, the prediction that *if not* statements should largely be endorsed appears to be upheld. The moderate rates of rejection we observe here can, as in Experiment 1, be attributed to implicature: for instance, some participants (≈20%) were inclined to reject *Some marbles have a dot if they are not red* when 40% of red and 40% of blue marbles had dots. We discuss the influence of implicatures on *if not* as well as *unless* statements in more detail below.

Moving to the *unless* data, the results support (30) over (29), repeated here, as a formulation of the AtB prohibition.

(29) $Q[C]M$ *unless* $R \Rightarrow Q[C-R]M \land \neg Q[C]\mathbf{M}$ **(Option 1)**

(30) $Q[C]M$ *unless* $R \Rightarrow Q[C-R]M \land \neg \mathbf{Q}[\mathbf{C} \cap \mathbf{R}]\mathbf{M}$ **(Option 2)**

Consider the *some* condition, for example (Figure 5, center panel). On (29), we expect *Some marbles have a dot unless they are blue* to be unsatisfiable: it requires simultaneously that some red marbles have a dot, and that no marbles have a dot. This prediction is clearly falsified by the very

high ($\approx$85%) endorsement of this sentence at proportion 0 in the middle panel of Figure 5—i.e., when 40% of red marbles have a dot and no blue marbles do. In contrast, (30) requires that some red marbles have a dot, and that no blue marbles do. This is consistent with participants' strong tendency to endorse the sentence at target proportion 0, and low rate of endorsement ($\approx$20%) at all higher target proportions.

The analysis of *most* and *few* is similar, though more complex due to the inherent vagueness of these quantifiers. In the case of *most*, (30) predicts that *Most marbles have a dot unless they are blue* should mean that most red marbles have a dot, and that it is not the case that most blue marbles have a dot. In line with this prediction, the acceptability of this sentence is high when the proportion of blue marbles is low, and declines steadily as the proportion increases (Figure 5, left panel). This analysis also explains why *Most marbles have a dot unless they are blue* is least acceptable when 80% of blue marbles have a dot: this is the same point at which *Most blue marbles have a dot* is most acceptable in our control data (see Figure 4).

Along the same lines, *Few marbles have a dot unless they are blue* is most acceptable when the proportion of red marbles most clearly counts as 'few' (at 20%) and when the proportion of blue marbles most clearly does not count as 'few' ($>$50%) (Figure 5, right panel). The rate of endorsement of this sentence declines to a low value ($\approx$20%) at the point where 20% of blue marbles have a dot. This is the point at which the description *Few blue marbles have a dot* would be most appropriate, according to our control data. Interestingly, *Few marbles have a dot unless they are blue* was endorsed at a slightly higher rate when no blue marbles have a dot. This likely reflects the fact that the description *Few blue marbles have a dot* is not very natural in this condition (endorsed by only $\approx$40% of participants; Figure 4). While we will not pursue the issue further, this result suggests that the AtB prohibition may sometimes take into account the implicature from *few* to *not none*, even though the relevant instance of *few* occurs in a deeply embedded and not-at-issue context. (A similar effect, presumably associated with the implicature from *some* to *not all*, is observed in the *some* data at proportion 1.)

There was one notable difference between Experiments 1 and 2. When the AtB condition associated with *unless* was falsified—which we have tentatively classified as triggering a presupposition failure—the rates of endorsement in Experiment 2 were noticeably higher than in Experiment 1. In the original experiment, the average endorsement rate in such cases was about 10% (min=5.5%, max=14.7%). In Experiment 2, the average endorsement rate was about 25% (min=14.7%, max=46.8%). Some of this variability may be attributable to uncertainty about the truth-conditions of the vague quantifiers used in Experiment 2. For instance, the maximum value of 46.8% occurred when 40% of the target marbles had dots. It may be that many subjects simply did not consider *Few target marbles have dots* to hold in this case, leading to mixed results as to whether the AtB prohibition was violated.

Overall, the results of Experiment 2 support the existence of a prohibition on Across-the-Board contexts, with the specific formulation given in (30). The results are also consistent with our suggestion that the second conjunct, the **AtB prohibition**, is a lexical presupposition of *unless*.

5.6 Discussion: Implicatures

Next, we consider the implications of the new data for the pragmatic inferences that – according to our analysis of Experiment 1 – affect *unless* and *if not*. Recall that the entailments in (48a) and the biconditionality implicatures in (48c) are postulated to affect both *unless*- and *if not*-conditionals, while we have proposed that the AtB prohibition in (48b) is a presupposition associated only with *unless*.

(48)   Most/Some/Few marbles have a dot $\begin{cases} \text{if they are not blue.} \\ \text{unless they are blue.} \end{cases}$

     a.   *if not*   Most/Some/Few red marbles have a dot.

                                              entailed by: *if not, unless*

     b.   *AtB prohibition*   It is not the case that most/some/few blue marbles have a dot.

                                              presupposed by: *unless*

     c.   *not if*   Most/Some/Few blue marbles do not have a dot.

                                              implicated by: *if not, unless*

As explained in §4.2, we expect to see reduced acceptance rates for both conditional types when the experimental context does not support biconditionality.

With *if not*, this prediction is upheld under the quantifiers *most* and *few*. Table 8 shows the average endorsement rates for these conditions, in conditions where the proportion of target colour marbles with dots was greater/less than 0.5. For *most-if not*, the >0.5 conditions fail the biconditionality condition (48c), since more than half of the blue marbles have dots. The <0.5 conditions plausibly support biconditionality. Thus, the fact that *most-if not* statements have markedly lower acceptance in the former condition (>0.5) than in the latter (<0.5) suggests an implicature to biconditionality. These results are mirrored in the data from *few-if not*, as predicted: >0.5 conditions support biconditionality, and the <0.5 conditions do not.

Target props.	MOST	FEW
> 0.5	63.1	**83.4**
< 0.5	**89.5**	72.9

Table 8: Average endorsement rates for *if not* in *mid* non-target conditions

The *some-if not* condition is interestingly different. Here, only one test condition for *some-if not* fails to support *not if*/biconditionality ("Some blue marbles do not have a dot"): the target dot proportion 1, where all of the blue/target colour marbles have dots. Thus, if the implicature in (48c) were the only pragmatic factor affecting the use of *if not*, we would expect near-ceiling endorsement with *some* in all other target dot conditions (0-0.8). This is not the pattern we observe: instead, endorsement is near ceiling only when at proportion is 0, and is markedly lower across all other proportions.

This pattern suggests that some other pragmatic inference is at play. Our results suggest instead an implicature that *none* of the target colour marbles have dots—in other words, an implicature with the same form as the AtB prohibition (here, "It is not the case that some blue marbles have dots"). This inference is only supported at target dot proportion 0, and as a result there is reduced acceptance in the *some-if not* condition at proportions 0.2-1. The fact that endorsement is still fairly high (between 70-80%) suggests that the effect is indeed due to a failed implicature, rather than a presupposition or entailment. As we saw in Experiment 1, false items were rejected almost unanimously. For a failed presupposition, we only need to look at the non-AtB proportions 0.2-1 in the *some-unless* condition, where acceptance was averaged around 20%.

The existence of an anti-AtB implicature (48b) under *if not*, separate from the implicature to biconditionality (48c), is only observable among our experimental conditions with *some*. This is because the implicature is strictly weaker than, and so masked by, the biconditionality inference with the other quantifiers used in our experiments. In contrast, with *some* the anti-AtB implicature is stronger, and so appears clearly in the data. The results of Experiment 2 thus suggest that *if not* generates both implicatures. In the next section we will consider the anti-AtB condition in more detail and connect it to the *conditional strengthening* phenomenon discussed in previous literature.

Lastly, we examine the Experiment 2 evidence of implicatures associated with *unless*. These data are consistent with the hypothesized effect of a *not if* implicature as stated in (48c). For *most-unless*, we observe a decline in endorsement rates as the target dot proportion goes from 0 to 1. This patterns with the decline in endorsement rates observed for *most-if not*, but has a markedly steeper slope. Since low endorsement rates are predicted independently for conditions that fail to satisfy the AtB prohibition (i.e. 0.6-1 for *most-unless*), the effect of presupposition failure is also expected to result in a steeper slope for *unless* than for *if not*.

As with *if not*, the *few-unless* data mirror the *most-unless* data. Endorsement rates steadily increase from target dot proportion 0 to 1, starting at a lower endorsement rate, and increasing with a steeper slope than *few-if not*. Presupposition failure here is expected at target dot proportions 0.4 and below, again driving low acceptance rates on these conditions. Since, as discussed above, the AtB prohibition is stronger than if *not if* inference for *some*, we do not expect to see anything other than the effect of presupposition failure in the data from this quantifier.

6 A revised semantics and pragmatics for *unless*

Here, we expand on the experimentally-driven proposals made in §4-§5. We present a new, more formal account of the semantics and pragmatics of *unless*, with implications for the theory of conditional statements more generally. We draw on the account in Nadathur (2014), but go beyond the earlier work.

Returning to the schema outlined at the beginning of §5 (example 42), the following three inferences appear to be associated with *unless* and *if not*:

(49) $Q[C]M$ COND R, where COND $\in \{if\ not, unless\}$
    a. *if not* $Q[C \wedge \neg R]M$
    b. *AtB prohibition* $\neg Q[C \wedge R]M$
    c. *not if* $Q[C \wedge R]\neg M$

Our current understanding of (49) is as follows.

First, we have seen no evidence against treating the *if not* inference as an entailment of *unless*. We therefore take (49a) to belong to the asserted content of an *unless* statement, as well as of an *if not* statement. We claim, additionally, that this is the full assertive content of *unless*; other associated inferences represent presuppositions or implicatures.

Second, in §4 we suggested that the AtB prohibition (49b) is a lexical presupposition of *unless*. Based on the results of Experiment 2, we have also proposed that *if not* has an anti-AtB implicature. Thus, (49b) is relevant to the interpretation of both items, but non-conventionally and more weakly in the case of *if not*.

Third, §4 suggests that both conditionals implicate the *not if* inference (49c). The Experiment 2 data, while consistent with this hypothesis, did not shed new light on the matter. More specifically, Puzzle (C)—the asymmetry between *unless* under *every* vs. *no*, discussed in §4.3—remains unresolved. We have clear evidence from Experiment 1 that an inference to the *not if* conditional (or to biconditionality in general) arises more strongly with *unless* composed with *every* than it does for either *if not* or for *unless* composed with *no*. Results from Experiment 2 are consistent with a view on which biconditional inferences for *unless* are stronger than those for *if not*. In addition, we have tentatively suggested that this strength differential (*unless* vs. *if not*) is reduced under composition with any negative quantifier (e.g., *few*), and not only under *no*. Although we believe that the data from Experiment 2 is inconclusive on this last point, we explore it in further detail in what follows.

We believe, however, that the evidence for treating *unless*-biconditionality pragmatically, rather than truth-conditionally, is conclusive. This follows from our experimental data, as well as from the arguments presented in §2.3.2. Consequently, while we remain in need of a pragmatic explanation for the strength differentials comprising Puzzle (C), apparently variable biconditionality does not pose a compositionality problem for the interpretation of *unless*.

Our investigation provides no evidence for any inferences associated with *if not* and *unless* beyond those in (49). While the two conditional forms apparently have similarities beyond their shared assertive content (49a), we remain committed—as per our experimental results—to the existence of important differences in the communicative potential of *if not* vs. *unless*. We believe these difference have consequences for speaker choice between the two alternatives. We will propose that these differences follow from the contrast in the inference type of the AtB prohibition as associated with *if not* vs. *unless*, and from the way in which the AtB prohibition interacts with conditional perfection (i.e., the *not if*/biconditionality implicature).

We first describe a semantics for *unless*, drawing on Leslie's (2009) 'modalized restrictor' conditional. This captures the asserted content (49a), while circumventing the embedded-conditional issues raised by Higginbotham (1986). We then suggest that the AtB prohibition as formulated in (49b) represents a special (possibly default) case of a general appropriateness condition for conditional statements. This condition functions as an implicature for *if not*, but as a presupposition for *unless*, accounting for the effects observed in across-the-board conditions in Experiments 1-2. Finally, we argue that the implicated *not if* content (49c) is driven by the same considerations as conditional perfection for bare (unquantified) conditionals, and suggest that this, plus the complexity added by composing a conditional with quantifiers of varying types, account for the form of the *not*

if inference, as well as for its apparently variable effects. While we cannot at this time provide a complete analysis of Puzzle (C), we sketch some lines of inquiry that may be fruitful to pursue.

6.1 The semantics of *if not* and *unless*

The proper semantic treatment for quantified indicative conditionals like (50a)-(50b) is a longstanding problem (Higginbotham 1986).

(50)   a. Everyone will fail if she does not work hard.

       b. No one will succeed if she does not work hard.

On the assumption that succeeding is the opposite of failing, (50a) and (50b) have the same intuitive meaning. The problem lies in deriving this equivalence compositionally – that is, in assigning a single semantic representation to *if* which will produce the same truth conditions for (50b) as for (50a). Higginbotham (see §2.1) discusses the inadequacy of embedding material *if* under the nominal quantifiers *every* and *no*, but the problem goes beyond this. Even on the received 'restrictor' analysis, embedding an *if*-clause under the nominal quantifier fails to derive the desired equivalence.[10]

     A number of solutions have been proposed for the quantified conditional problem (see von Fintel 1998, Leslie 2009, Kratzer in press). Here we adopt Leslie's 'modalized restrictor' conditional, which builds on Kratzer's (1986) influential treatment.[11] Kratzer proposes that unquantified conditional statements contain a covert universal modal quantifier (*must*), on which an *if*-clause operates as a domain restrictor. In (51), *Acc* picks out the set of worlds relevant for evaluation (according to modal flavour).

(51) q *if not* $p := \forall w[w \in Acc \wedge \neg p(w)]q(w)$ ($\leftrightarrow$ *must q if not p*)

Leslie extends the restrictor analysis in two ways. First, she adopts a proposal from von Fintel (1998), on which *if*-clauses can act as restrictors of nominal quantifiers, as well as of modal quantifiers. This produces reasonable and equivalent truth conditions for (50a)-(50b).

(52)   a. Everyone will fail if she does not work hard.
         $= \forall x[\neg \text{work}(x)]\text{fail}(x)$
         All individuals who do not goof off succeed.

---

10 This structure yields the truth conditions below. Crucially, while (ii-a) is a reasonable interpretation of (50a), (ii-b) is too weak: not working hard should not simply make failure possible, but instead guarantee it.

     (ii)   a. Everyone will fail if she does not work hard.
            $= \forall x(\forall w[w \in Acc \wedge \neg \text{work}(x,w)]\text{fail}(x,w))$
            All individuals are such that all non-hard-working worlds are failure-worlds.

          b. No one will succeed if she does not work hard.
            $= \neg \exists x(\forall w[w \in Acc \wedge \neg \text{work}(x,w)]\text{succ}(x,w))$
            $= \forall x \neg(\forall w[w \in Acc \wedge \neg \text{work}(x,w)]\text{succ}(x,w))$
            $= \forall x(\exists w[w \in Acc \wedge \neg \text{work}(x,w)]\neg \text{succ}(x,w))$
            All individuals are such that (at least) one of their non-hard-work worlds is not a success-world.

11 See also von Fintel & Iatridou (2002), Higginbotham (2003), Huitink (2009).

b. No one will succeed if she does not work hard.
$= \neg\exists x[\neg\text{work}(x)]\text{succ}(x)$
$= \forall x[\neg\text{work}(x)]\neg\text{succ}(x)$
All individuals who do not work hard do not succeed.

This move comes at a cost, which Leslie illustrates as follows: suppose the domain of quantification for (50b) is a class of students including Meadow, who is so bright that she will succeed regardless of the amount of work she puts in. Intuitively, Meadow ought to falsify (50b), since she will succeed even if she does not work hard. Crucially, the truth of (50b) should not depend on what Meadow chooses to do in the actual world. (52b) is unable to capture this intuition, illustrating that our evaluation of quantified conditionals like (50a)-(50b) depends not only on what happens in the actual world, but also on alternative possibilities. In other words, the modality introduced for the unquantified conditional in (51) remains relevant even when *if* restricts a nominal quantifier.

Leslie proposes to 'modalize' the extended restrictor analysis (52). Specifically, she suggests that, like their unquantified cousins, quantified indicative conditionals contain a covert universal modal. In cases like (50a)-(50b), the modal takes wide scope, producing the interpretations in (53).

(53) a. Everyone will fail if she does not work hard.
$= \forall w[w \in Acc](\forall x[\neg\text{work}(x,w)]\text{fail}(x,w))$
All relevant words are such that everyone who does not goof off succeeds.

b. No one will succeed if she does not work hard.
$= \forall w[w \in Acc](\neg\exists x[\neg\text{work}(x,w)]\text{succ}(x,w))$
$= \forall w[w \in Acc](\forall x[\neg\text{work}(x,w)]\neg\text{succ}(x,w))$
All relevant words are such that everyone who does not work hard does not succeed.

Leslie deals only with cases where a conditional statement contains at most one overt quantifier. This is also sufficient for present purposes. We note, however, that a potential benefit of the modalized restrictor account is that it easily extends to conditional statements involving overt realizations of both nominal and modal quantifiers (see Geurts 2004: for further discussion).

Since (as we have argued) *unless*-conditionals share the asserted content of *if not*-conditionals, the counterparts of (53) with *unless* have essentially the same analysis. The primary difference is that, with *unless*, negation is specified in the lexical entry rather than appearing as a separate element in the syntax. Leslie's proposal thus provides a working semantics for both our conditional types. Crucially, this analysis assigns the same semantic representation to *if not* and *unless* regardless of quantificational context, avoiding Higginbotham's problem.

6.2 Conditional strengthening, across-the-board contexts, and presupposition

Our experiments show that there are empirical differences in the interpretation of *unless* vs. *if not* which cannot be motivated by shared assertive content. Here, we propose that the core distinction lies in content that is carried by *unless* as presupposition, but which associates with *if not* as implicature. This content is responsible for the anti-AtB effects seen in Experiments 1-2; we will appeal to the difference in inference type to account for categorical differences between *unless* and *if not*, where they arise.

Experiment 1 motivated the AtB prohibition (49b) as a precondition for the use of *unless*. Experiment 2 provided evidence for an implicature of the same form associated with *if not*. We believe that the anti-AtB inference arises as a special case of an inference which conditionals in general invite about their utterance contexts. We first discuss the properties of a generalized **conditional strengthening** implicature, before dealing with the consequences of its lexical status for *unless*.

In connection with *conditional perfection*, it has been observed that a conditional *if p, then q* suggests the possibility that its consequent is not realized (see Zaefferer 1991, von Fintel 2001, Franke 2009).

(54)   Robin will be on time if Sanya is coming.
      ⤳ *It is not necessarily (unconditionally) the case that Robin will be on time.*

In other words, a conditional like (54) invites an inference to the negation of its consequent.[12] Adopting von Fintel's (2001) terminology, we call this inference **conditional strengthening**. At first pass, we might model it as a standard scalar implicature:

(55)   **Conditional strengthening:** given a conditional operator COND, and two propositions *p* and *q*, the statement *q* COND *p* invites an inference that the speaker is unwilling or unable to commit to the bare consequent *q*.

Assuming that a conditional's consequent represents a salient pragmatic alternative to the full conditional, the consequent will rank as more informative and thus stronger on a Horn scale. Use of the weaker alternative in this situation typically signals the speaker's unwillingness to commit to the stronger option. One reason for this might be that she lacks evidence to support the truth of *q*, and thus chooses the conditional in accordance with Grice's Quality maxim. However, if we have reason to think she is well-informed about the truth value of *q*, we might instead infer that she has evidence against *q*—in other words, that she knows *q* to be false in relevant situations. When combined with the assertion *q if p*, this produces the familiar 'not across the board' inference (49b).

(56)   Robin will be on time if Sanya is coming.
      a. *Assertion:* All worlds where Sanya comes are worlds where Robin is on time.
      b. *Non-epistemic conditional strengthening:* ⤳ There are accessible worlds in which Robin is not on time.
      c. (56a) & (56b) ⊨ Not all worlds where Sanya does not come are worlds where Robin is on time.

---

12 On the restrictor analysis in (51), the inference represents the negation of the full consequent, including a covert universal modal. Moving to the extended restrictor analysis, then, we predict—intuitively correctly—that a parallel inference invited by a nominally-quantified conditional will represent the negation of the nominally-quantified consequent. This generates version 1 of the AtB prohibition, in (29)—e.g., *Every marble has a dot if it is not blue* invites an inference to *It is not the case that every marble has a dot.*

The stronger, non-epistemic implicature is presumably the relevant one for understanding our experimental data, since we provided a fully-specified visual display. With nominally-quantified conditionals this implicature combines with the assertive content to generate, for instance:[13]

(57) Every marble has a dot if it is not blue.
    a. *Assertion:* All non-blue marbles have dots.
    b. *Non-epistemic conditional strengthening:* ⤳ Not every marble has a dot.
    c. (57a) & (57b) ⊨ Not all blue marbles have dots.

This reasoning is complicated by the results of the *some-if not* condition in Experiment 2. There we observed the effects of an anti-AtB implicature, but it cannot be understood in terms of the conditional strengthening mechanism just mentioned: *Some marbles have a dot if they are not blue* is inconsistent with the negation of its consequent *Some marbles have a dot*.

Alternatively, we could focus on the fact that the consequent is less structurally complex than the full conditional. Assuming a *ceteris paribus* preference for less complex utterances, the choice of a dispreferred alternative should implicate—in line with Grice's Manner maxim—that the speaker has some reason for choosing the conditional over its consequent. This reason cannot be that the speaker lacks sufficient evidence to assert the (entailed) consequent, so there must be some other axis along which it is preferable. In this case it is not difficult to identify the reason for the preference: the conditional is strictly more informative, and so preferred by Quantity reasoning. However, the example reveals that conditional strengthening effects cannot always be reduced to Quantity reasoning. If these effects are indeed fully general, they must rely in the first place on some type of Manner reasoning.

Generalizing across the three cases in (56), (57), and from *some*-conditionals, we suggest the following revised characterization of conditional strengthening.

(58) **Conditional strengthening (revised):** A speaker who utters a conditional statement *q* COND *p* has a reason for avoiding an assertion of the bare consequent *q*.

Broadly, we are suggesting that the use of a conditional invariably leads to pragmatic reasoning in which the conditional is contrasted with the simpler alternative offered by the bare consequent, which does not stand in any necessary relation of informativity to the full conditional.

On this view, conditional strengthening is a special case of the general class of **Need-a-Reason (NaR) implicatures** identified by Lauer (2013). NaR implicatures belong to the Gricean tradition, but arise on the basis of a strong communicative preference for simpler utterances. Lauer motivates these inferences with the case of the *ignorance implicature* that often arises with disjunctions.

(59) John is in Paris or he is in London.
    ⤳ The speaker is unable/unwilling to say which.

In (59), either disjunct (*John is in Paris*, *John is in London*) represents an alternative to the full utterance; moreover, both alternatives are simpler than the full disjunction in terms of both processing and comprehension costs.

---

13 We are assuming here that the specific reference induced by the display/description pair preclude the inclusion of a wide-scope modal for our target conditionals.

Lauer argues that, as with other Gricean inferences, violating the preference for simpler utterances indicates that it is being overridden by some other communicative need: i.e., there must be some justification for the speaker's decision to use a more complex form. In the case of (59), this reason cannot be that the full utterance is more informative, since either alternative is actually informationally stronger than their disjunction. Consequently, some other reason must be inferred. If this is taken to relate to the speaker's epistemic state, the familiar ignorance implicature arises.

Conditional strengthening has a similar analysis: the more complex conditional invites an inference about the speaker's justification for choosing it over the simpler bare consequent. In some cases (e.g., the *some-if not* examples) the greater informativity of the conditional is sufficient reason. In other cases (e.g., *most-if not*), another reason must be operative—for instance, the speaker might avoid the bare consequent because it is false. In these cases, our predictions align with von Fintel's original analysis in (55).

For *unless*-conditionals, we propose that the categorical AtB effects observed in our experiments arise from a conventional association with the same conditional strengthening condition, in this case as a presupposition. This predicts an important difference between *if not* and *unless*. If conditional strengthening is conventionally associated with *unless*, a speaker's choice to use this item indicates that the speaker is aware of the implication and willing to take responsibility for it. While a Gricean implicature could arise inadvertently with an *if (not)* conditional, the choice of *unless* requires the speaker to have a pre-utterance awareness of the justification. The following minimal pair illustrates.

(60) a. Robin will leave if Sanya does not call her. (?)Actually, she's going to leave no matter what.

   b. Robin will leave unless Sanya calls her. ??/#Actually, she's going to leave no matter what.

The second sentence of (60a) reads as an elaboration to correct an unwanted implicature. In contrast, (60b) feels like a contradiction. Our suggestion is that the choice of *unless* directly conveys the speaker's belief that she and her interlocutors are in a context that would not support the use of the bare consequent. Immediately denying this presupposition in (60b) thus produces a stronger sense of infelicity than the denial of an otherwise identical implicature in (60a), where the speaker can with some effort distance herself from the strengthening inference.

This clash explains why, in the context of Experiment 1, *unless* statements appeared subject to a categorical prohibition on AtB contexts. The experimental context simply provided no recoverable justification for a conditional description. In these contexts, use of *unless* violated the NaR precondition: these statements were thus penalized (85.3% rejection for *every-unless* at target dot proportion 1, 94.5% rejection for *no-unless* at 0) as representing 'unjustified' and therefore infelicitous/incorrect uses of a conditional form.

While Lauer (2013) predicts NaR implicatures to be non-optional,[14] we note that failed implicatures come apart from failed presuppositions with respect to whether or not their content can be

_____

14 Non-optionality is a particularly interesting feature of NaR implicatures, setting them apart from more standard Gricean inferences. Non-optionality follows from the idea that the low-cost preference is in some sense paramount—thus, any deviation from the predictions of this preference is *only* possible on the basis of some other communicative need. That is, when we choose to use a more complex alternative, there is necessarily a reason for this. The resolution of the NaR inference will differ based on discourse context and/or world knowledge, but the core inference itself remains.

said to be directly conveyed. This is the reason for the distinction in rejection rates for *unless* and *if not* in AtB contexts. *If not* is rejected with lower frequency than *unless* in the same contexts (33.3% rejection for *every-if not* at proportion 1, 40.0% for *no-if not* at 0), but with higher frequency than *if not* statements with failed biconditionality implicatures in the middle range of target dot proportions (avg. 21.0% rejection for *every-if not*, 21.6% for *no-if not* over proportions 0.2-0.8). Our data are consistent with the idea that participants struggle to ignore an anti-AtB inference when making truth value judgements, but were more easily able to set aside *post hoc* pragmatic reasoning than they were able to absolve the speaker of responsibility for presupposing a justification. In other words, *unless* statements were penalized more severely because the speaker is held responsible for the conditional strengthening condition.

This analysis also provides us with a way of making sense of one lingering puzzle from Experiment 2. As discussed in §5.5, rejection rates in some plausibly AtB contexts were less catgorical for *unless* statements modifying *most, some*, and *few* than they were for *unless* modifying universal quantifiers in Experiment 1. We suggest that this effect can be attributed to the vagueness of *most* and *few*, coupled with the inherent flexibility of NaR reasoning. Some participants may have reasoned as follows: a possible justification for choosing the *unless* statement rather than its consequent is that there was a salient difference between the dot proportions in the target vs. non-target colors, and the quantifier clearly held of the target color and less clearly of the non-target. For example, we might expect higher acceptance rates for (61) when 80% of non-target marbles have dots, and only 60% of target marbles have dots, as compared to a case where 80% of target marbles have dots as well.

(61)   Most marbles have a dot unless they are blue.

     a.   Most non-blue marbles have a dot. (80%: Clearly true)

     b.   Most marbles have a dot. (Only 60% of blue: Less clearly true)

Here, (61) might be assessed by comparing the felicity of (61a) to (61b). If (61a) is more felicitous than (61b) due to the vagueness of *most*, some participants may have interpreted this contrast as enough to satisfy the conditional strengthening precondition would be satisfied. This is generally consistent with our experimental results (see also Appendix A).

Cast as a presupposition, conditional strengthening seems at first rather atypical, since it involves a certain amount of contextual input to be fully resolved. We observe, however, that although the required information is of a different sort, this is not dissimilar from anaphora resolution. In the latter case, speaker presuppositions of existence, uniqueness, etc, are satisfied by drawing on discourse context and local information; the same is true for conditional strengthening. To draw out the parallel farther, an anaphor can produce infelicity if its presupposition of a uniquely identifiable reference is not easily satisfied by reference to context: similarly, *unless* produces infelicity if the speaker's reasoning (in this case about the markedness of the excepted/target set) is not recoverable. In both cases, the speaker can supplement the presupposition-carrying item with additional information that aids in resolution: for *unless*, this might be supplied by an *in which case* continuation, as in (62).

(62)   Robin will leave unless Sanya calls her, in which case she might stay longer.

We suspect that such continuations will primarily be used in 'questionably' exceptive cases – for instance, where the speaker singles out the target set not because it fails to satisfy the main predicate, but because it does so in some marked or non-canonical way.

In sum, treating conditional strengthening as an implicature for *if not* but as a presupposition for *unless* leads us to draw the following distinction between the two conditionals. Both *if not* and *unless* generate an inference that the speaker has a reason to avoid asserting *q*. However, *unless* is stronger insofar as it expresses the speaker's active belief that this reason is recoverable from context. In a sense, this presupposition captures the original idea behind the exceptive account (see, e.g., Dancygier 1985): *unless*-statements differ from their *if not* counterparts in that they conventionally *draw attention* to the relationship between their "main generalization" (the consequent) and the excepted set. Put another way, when a speaker indicates that the excepted set is not relevant for consideration of the main generalization in an *if not* statement, a listener may be inclined to wonder why the speaker chose to exclude the excepted set. In contrast, an *unless*-statement actively communicates that there is something divergent about the excepted set, and, as a result, checking that the presupposition holds requires active attention to the excepted set.

6.3   Conditional perfection and quantified generalizations

An important outstanding piece of the puzzle is the tendency of *unless*-statements to be interpreted biconditionally. This was the presumable source of the lowered acceptability of *unless*-statements relative to matched *if not*-statements in non-biconditional contexts—e.g., the middle range of proportions for statements under *every*. We have argued that this pattern is related to *conditional perfection*. Here, we expand on this, offering an explanation of why *unless* seems to perfect more strongly than *if not* in many contexts.

6.3.1   A sketch of conditional perfection

There is an extensive literature on conditional perfection with a variety of theoretical accounts (see van der Auwera 1997, Horn 2000). The account sketched here is based on Franke (2009) and Nadathur (2013), who argue that conditionals tend to be perfected when they are interpreted as providing (complete) answers to polar questions on their consequents. We will illustrate the reasoning without attempting to provide a detailed explanation and defense: see the original sources for more detail and explanation of apparent counter-examples.

Consider first a case where the question under discussion (QUD; Roberts 2012) is established overtly:

(63)   A: Will Robin come to the party?
       B: If you don't invite Xavier, Robin will come.
          ⤳ . . . *and if you do invite Xavier, Robin won't come.*

Here, B's conditional response suggests not only that Robin will attend if Xavier is not invited (the literal content of (63)B), but also that Robin will *not* attend if Xavier is invited. Taken together, this produces the perfected interpretation, on which Xavier's invitational status fully determines whether or not Robin comes to the party.

The QUD-based story goes roughly as follows. A's question sets up a context in which *B* is presumed to be informed about the value of $?\phi = $ *Will Robin come?*. However, the literal content of B's response only settles $?\phi$ in a subset of these futures. If *B* is indeed well-informed and is attempting to give a full answer, we must infer some additional content that addresses what happens in the possible futures in which we do invite Xavier, filling in the gap between the literal meaning and a complete resolution of the QUD. One way to add this content would be to suppose that the consequent is true no matter what. However, this enrichment would run afoul of the conditional strengthening inference discussed above: if the speaker's information had supported *Robin will come*, she presumably would have said this rather than using the more prolix and less informative conditional construction. The only way to enrich B's response into a complete answer without running afoul of conditional strengthening is to infer instead that the consequent is false across Xavier-inviting futures. This inference, combined with the literal meaning of the conditional, yields the biconditional interpretation.

In this sketch of a conditional perfection inference for an *if not* conditional responding to an explicit polar question, it is not just the QUD that is central to the inference, but also crucial assumptions about the speaker, context, and conversational goals. Consequently, we expect perfected interpretations of the *if not* conditionals in contexts which (a) make the consequent's resolution of primary salience to the conversation, and (b) support assumptions of speaker epistemic authority and responsibility, regardless of whether such contexts involve an overt polar question or not.

Using the same reasoning, we expect *unless* conditionals to give rise to perfected interpretations under the same conditions.

(64)   A: Will Robin come to the party?
       B: Robin will come unless you invite Xavier.
          ⤳ *Robin will not come if you invite Xavier.*

6.3.2   Speaker authority and responsibility with *unless*

On our analysis, the only meaning difference between *if not-* and *unless*-conditionals is that the conditional strengthening NaR inference is a conventional presupposition of the latter, rather than being a weaker inference generated by context-sensitive pragmatic reasoning. Our suggestion is that the effect of presuppositional conditional strengthening duplicates some of the work done by a polar question on the conditional consequent in (63), thus facilitating a perfection inference for *unless* even in the absence of the right kind of QUD.

A speaker who chooses a lexical item or construction that carries a presupposition (especially as opposed to an assertorically equivalent alternative which lacks the same presupposition) may, in general, be taken to indicate her belief that the utterance context supports her choice. Thus a speaker who chooses *unless* communicates a belief that the utterance context supplies a reason for avoiding the unconditional consequent, taking direct responsibility for the existence of this reason.

In some cases a reason for avoiding the bare consequent is readily recoverable from the common ground. Given an *unless* conditional of the form $Q[C]M$ UNLESS R, it may be clear from the context that some part of the excepted set R falsifies M, or, alternatively, it may be established that the speaker's knowledge is restricted to the part of C which is not also R. Consider, for instance, a

variant of our experimental design in which it is known that, while the speaker is aware that the box contains both red and blue marbles, she has only seen the red ones. In this context, a listener will easily recover epistemic uncertainty as the felicitous precondition justifying the use of *unless* in the statement "Every marble has a dot unless it is blue" — which could easily be followed by "... in which case I have no idea." In this type of context, any inferences drawn over and above conditional strengthening will rely heavily on contextual details.

If a reason for avoiding the bare consequent is not readily recoverable, the listener must accommodate the speaker's presupposition that there is a reason. By making such a presupposition, the speaker is both assuming responsibility for its truth and directing the listener's attention to the excepted set, where the reason must be recovered. The attentional effect of the choice of *unless* can be seen in the way it constrains the interpretation of subsequent propositional anaphors differently from matched *if not*-conditionals.

(65) a. Robin will come to the party if you don't invite Xavier. In that case, she'll stay late and have a great time.

   b. Robin will come to the party if you don't invite Xavier. # In that case, she'll remain at home.

In the *if not* examples, the attentional focus remains on the set explicitly mentioned in the antecedent—the worlds in which you *don't* invite Xavier—rendering the follow-up in (65a) sensible while (65b) is contradictory. In contrast, the matched *unless* examples redirect attention to the worlds in which you *do* invite Xavier, reversing the felicity of the examples.

(66) a. Robin will come to the party unless you invite Xavier. In that case, she'll remain at home.

   b. Robin will come to the party unless you invite Xavier. # In that case, she'll stay late and have a great time.

In sum, we are suggesting that *unless*-conditionals favour conditional perfection more robustly and reliably than matched *if not*-conditionals because the *unless*-conditionals invariably draw attention, via their presupposition, to the status of the excepted set with respect to the complement's truth. The status of this set is, in turn, a crucial issue in the derivation of the conditional perfection inference. *If not*-conditionals also draw attention to this set in many contexts, but not as part of their conventional meanings, and any contextual support for these effects contribute additionally to the same pragmatic effect for *unless*-conditionals in matched contexts. *Unless*-conditionals lend themselves to perfection, therefore, even in the absence of a polar QUD, and we expect them to be perfected at higher rates as a result. This may account for the systematic, but non-categorical, divergences that we observed between *if not-* and *unless*-conditionals in our experiments.

6.3.3 The remaining puzzle

Both our experimental results and the preceding theoretical discussion support a picture in which *unless*-conditionals are, for subtle but principled pragmatic reasons, associated with biconditional inferences in a more robust and less context-sensitive way then their *if not*-counterparts. In light

of this conclusion, Puzzle (C) appears in a slightly different light: what is it about the quantifier *no* that neutralizes the attentional effects associated with a speaker's choice of *unless*? An initially plausible idea is that it is somehow related to the downward monotonicity of *no*, in contrast to *every*. However, this suggestion is immediately falsified by *few*, which is also downward monotone but showed in Experiment 2 the now-familiar pattern of *unless*-facilitated biconditional inferences. In this section we will discuss some additional connections involving attention the interpretation of anaphors and definites. While no firm conclusion will emerge, we think that the solution to this puzzle is likely to be found in subtle attentional effects along these lines.

Given the discussion in the previous section, the most straightforward approach would be to look for an effect of quantifier choice on the attentional split between *unless* and *if not*, as diagnosed by propositional anaphors. Unfortunately, the resolution of propositional anaphors does not appear to be affected by the choice between *every* and *no*.

(67)   a. Every student will succeed unless she skips class, in which case she'll be unhappy/#happy with her grade.

   b. Every student will succeed if she doesn't skip class, in which case she'll be happy/#unhappy with her grade.

(68)   a. No student will succeed unless she attends class, in which case she'll be happy/#unhappy with her grade.

   b. No student will succeed if she doesn't attend class, in which case she'll be un-happy/#happy with her grade.

This indicates that, according to this test, changing the quantifier from *every* to *no* is not enough to neutralize the attentional effects that we argued were responsible for the tendency of *unless* to generate stronger biconditional inferences.

The theory of strengthened biconditional inferences under *unless* offered above does not, therefore, extend directly to account for Puzzle (C). We do, however, wish to maintain that it may yet be related to complex and poorly understood pragmatic effects of quantifier choice on attention, as reflected in anaphora and the interpretation of definites. Tantalizingly similar phenomena seem to occur, for example, in complement anaphora, in which the resolution of a nominal anaphor in a following clause is affected by the choice of quantifier (Moxey & Sanford 1993, Nouwen 2003). Some quantified sentences $Q[R]M$, such as those with $Q =every$, *many*, or *some*, allow an immediately following anaphor to refer to the intersection of their restriction and nuclear scope ($R \cap M$, the "reference set"), while disallowing anaphora to the set difference of their restriction and nuclear scope ($R - M$, the "complement set").

(69)   Every/most/some senator(s) talked to Kennedy. They $\begin{cases} \text{enjoyed the conversation.} \\ \text{\# successfully avoided him.} \end{cases}$

   a. ✓ "The senators who talked to Kennedy ..."
   b. # "The senators who did not talk to Kennedy ..."

In contrast, negatively-oriented quantifiers like *few*, *no*, and *not many* prefer that the anaphor be resolved to the complement set.

(70) No/few/not many senators admire Kennedy. They $\begin{cases} \text{\# enjoyed the conversation.} \\ \text{successfully avoided him.} \end{cases}$

    a. # "The senators who talked to Kennedy ..."

    b. ✓ "The senators who did not talk to Kennedy ..."

Based on the quantifiers involved, monotonicity would seem to play a role in this generalization. Nouwen (2003) suggests that the connection with monotonicity is due to the fact that upward monotone quantifiers guarantee that the reference set is non-empty, while downward monotone quantifiers guarantee that the complement set is non-empty. This connection is highly suggestive, but there are several issues in drawing a precise analogy and we leave the possible connection as a further puzzle for future work.

A second possible connection is the interpretation of plural definites, which tend to be pragmatically resolved towards a universal interpretation when they appear in positively quantified or upward entailing contexts, but are usually interpreted existentially in negative or DE contexts (Schwarzschild 1991, Löbner 2000).

(71) a. Peter opened the windows.   (all of the windows)

    b. Peter didn't open the windows.   (any of the windows)

The pattern is familiar. Under the positive universal *every*, the set picked out by the complement of *unless* is inferred to universally fail the main generalization ("has a dot"), leading to reduced acceptability in the midrange. Under the negative universal *no*, the same set need not be homogeneous with respect to the main generalization, leading to improved acceptability in the midrange. The literature on plural definites has appealed to homogeneity presuppositions (Schwarzschild 1991, Szabolcsi & Haddican 2004), the strongest meaning hypothesis (Dalrymple et al. 1994), or to conversational goals of speakers and hearers (Malamud 2012) which tend to be correlated with the choice to use positive vs negative environments. We believe that these will be fruitful connections, but we refrain from exploring them in this space.

7   Conclusions and future directions

We have examined here in some detail a set of challenges for the existing exceptive solutions to the compositionality puzzle apparently posed by *unless*. These challenges, set out in §2, included, first of all, evidence from naturally-occurring data which suggests that *unless* is not semantically biconditional (contra von Fintel 1993, Leslie 2009); secondly, the apparent capability of *unless* to compose with non-universal quantifiers (e.g. *most, few*, contra von Fintel 1993); and, finally, the existence of strong intuitions (Figure 1) which suggest the relevance of an across-the-board condition. We presented experimental evidence, in two parts, investigating these challenges. Our first experiment provided clear data ruling out a representation of *unless* as semantically biconditional, and so demonstrated the need to reframe the original 'compositionality' problem as one of pragmatics, rather than semantics. While suggesting a close (semantic) similarity between *unless* and *if not*, Experiment 1 also verified intuitions about the relevance of a prohibition against the use of the former, but not the latter, in contexts where the predicate in the quantifier's scope failed to

distinguish (in any way) between the excepted and unexcepted sets. Experiment 2 clearly demonstrated the use of *unless* with the non-universal quantifiers *most, some* and *few*, and in addition suggested an interpretation of the across-the-board prohibition as the contextual realization of a broader requirement, presuppositional on *unless*, but arising as an implicature for *if not* conditionals.

On the basis of our experimental data, we motivated a number of revisions to the existing theoretical accounts of *unless*. On the whole, the most interesting aspects of *unless* are external to its semantic contribution, which we take to be equivalent to that of *if not* (on Leslie's modalized restrictor analysis for *if*-clauses). We locate the difference between *unless* and its close alternative *if not* in specific aspects of pragmatics: where *if not* conditionals are subject to a *conditional strengthening* implicature requiring the speaker to have reason for the use of a conditional form instead of the bare consequent, the requirement for such a reason is realized as a presupposition on the use of *unless*. This presupposition yields a stronger inference to *conditional perfection* in the case of *unless*, and thus accounts for (a) the prominence of biconditional representations of *unless* in the literature, (b) the reduced acceptability of *unless* as opposed to *if not* in most of our non-AtB contexts that did not support perfection. One of the most interesting empirical discoveries of the paper is the failure of *unless* to strengthen biconditional inferences under *no*. While we were not able to offer a theoretical explanation, we were able to confirm that it is a special pragmatic feature associated with *no* (as opposed to *every, many, few*, and *some*), and we suggested some lines for future inquiry that may help in understanding this new phenomenon.

The account of *unless* given in §6 leads to a number of avenues for further work. For example, our claim that conditional strengthening is a Need-A-Reason implicature in Lauer's (2013) sense leads to the prediction that the AtB effect should be mitigated in contexts which supply some non-truth-conditional reason for being especially interested in the restricted set—e.g., in the blue marbles in *Every marble has a dot unless it is red*. Since we invoked conditional strengthening in the derivation of biconditionality inferences, we would also expect that the strength of biconditionality inferences should be reduced in the same contexts.

Although we have argued against the exceptive treatments of *unless* offered in von Fintel (1993) and Leslie (2009), we ultimately believe that *unless* is closely related to exceptive phrases. In addition to the negative constraint, the literature on exceptive phrases deals with a number of conditions and restrictions which closely resemble those that are relevant to the interpretation of *unless* (Keenan & Stavi 1986, Hoeksema 1987, 1990, von Fintel 1993, Moltmann 1995, Garcia-Alvarez 2008). Indeed, an experimental investigation of exceptives similar to our Experiment 1 found a strikingly similar pattern of results, with *except* and *but* analogous to *unless* and negative restrictive relative clauses analogous to *if not* (Nadathur & Lassiter 2018).

This similarity even extended to the occurrence of a split under *every*, but not under *no*. Further comparative investigation of *unless* and exceptive phrases should be illuminating. Moreover, given the variety of exceptive phrases, and their differing syntactic behaviour—with respect to one another, as well as *unless*, an understanding of the extra-assertive and pragmatic features of exceptive phrases would provide an excellent testing ground for examining the broader questions about attention, presupposition, and pragmatic reasoning that this paper has only touched on.

Appendix: Additional data from Experiment 2

Our discussion of Experiment 2 examined only one of the three non-target dot conditions included in the experiment design. This appendix presents the remaining results, from the *high* and *low* non-target dot conditions for each of the quantifiers *most, some*, and *few*. Figure 6 shows the results from the *high* conditions for each quantifier (left=*most*, center=*some*, right=*few*), and Figure 7 shows the data from the *low* conditions. The numerical value of the non-target condition (which varies by quantifier) is included in the label for each graph. Error bars represent 95% confidence intervals. The numerical data corresponding to Figures 6-7 is given in Tables 9-10, respectively.
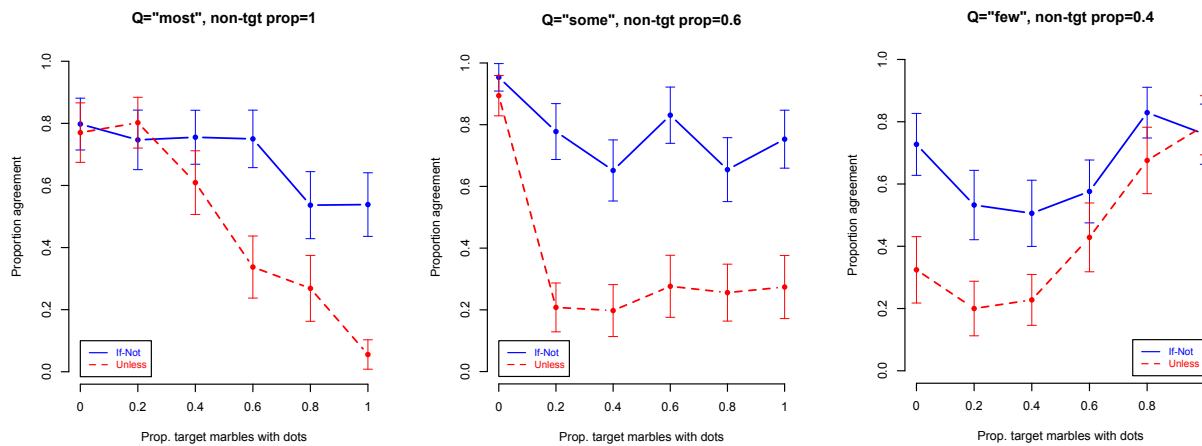

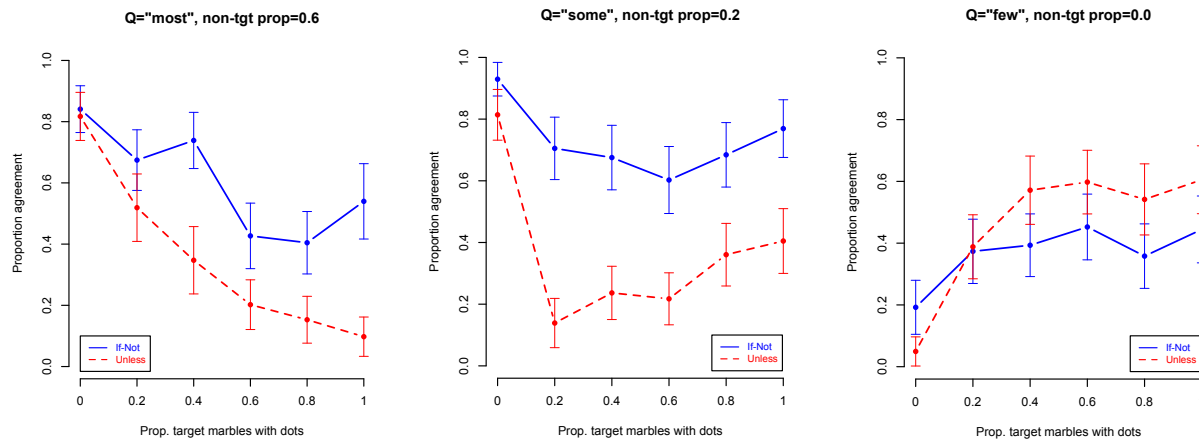Figure 6: Results from the *high* condition, by quantifier

Figure 7: Results from the *low* condition, by quantifier

		MOST		SOME		FEW	
		N	% agree	N	% agree	N	% agree
	0.0	89	79.8±8.3	86	95.4±4.4	77	72.7±10.0
	0.2	79	74.7±9.6	81	77.8±9.1	77	53.2±11.1
	0.4	94	75.5±8.7	89	65.2±9.9	85	50.6±10.6
If not	*0.6*	84	75.0±9.3	65	83.1±9.1	92	57.6±10.1
	0.8	82	53.6±10.8	81	65.4±10.4	82	82.9±8.1
	1.0	91	53.8±10.2	81	75.3±9.4	75	76.0±9.7
Prop. target marbles with dots	*0.0*	74	77.0±9.6	85	89.4±6.5	74	32.4±10.7
	0.2	91	80.2±8.2	101	20.8±7.9	80	20.0±8.8
	0.4	87	60.9±10.2	86	19.8±8.4	101	22.8±8.2
Unless	*0.6*	86	33.7±10.0	76	27.6±10.1	77	42.9±11.0
	0.8	67	26.9±10.6	86	25.6±9.2	74	67.6±10.7
	1.0	90	5.6±4.7	73	27.4±10.2	71	78.9±9.5

Table 9: Numerical data for *high* conditions, by quantifier

		MOST		SOME		FEW	
		N	% agree	N	% agree	N	% agree
	0.0	88	84.1±7.6	85	92.9±5.5	78	19.2±8.7
	0.2	86	67.4±9.9	78	70.5±10.1	83	37.3±10.4
	0.4	88	73.9±9.2	77	67.5±10.5	89	39.3±10.2
If not	*0.6*	82	42.7±10.7	78	60.3±10.9	84	45.2±10.6
	0.8	89	40.5±10.2	76	68.4±10.5	81	35.8±10.4
	1.0	63	54.0±12.3	78	76.9±9.3	81	44.4±10.8
Prop. target marbles with dots	*0.0*	93	81.7±7.9	86	81.4±8.2	81	4.9±4.7
	0.2	79	51.9±11.0	72	13.9±8.0	85	38.8±10.4
	0.4	72	34.7±11.0	93	23.7±8.6	77	57.1±11.0
Unless	*0.6*	94	20.2±8.1	92	21.7±8.4	87	59.8±10.3
	0.8	85	15.3±7.7	86	36.1±10.2	72	54.2±11.5
	1.0	82	9.8±6.4	84	40.5±10.5	76	60.5±11.0

Table 10: Numerical data for *low* conditions, by quantifier

We limit our discussion of these additional data to an assessment of their compatibility with the conclusions outlined in the discussion in §5.5, which was based on the *mid* non-target dot condition.

With the exception of data from the problematic *few-low* condition (see the discussion of control data in §5.5), results from the *high* and *low* non-target conditions parallel those from the *mid* condition. Endorsement rates in all experimental conditions reliably suggest that test *if not* sentences are true, *modulo* pragmatic factors. *Unless* acceptance rates in these conditions also replicate the earlier data in demonstrating the interpretability of *unless* composed with non-universal quantifiers.

The data in Figures 6-7 and Tables 9-10 also resemble the *mid* condition results with respect to the upper bound on endorsement rates on those test items (27.6%-33.7% on *high*, 20.2%-40.5%

on *low*, excluding *few-low*) predicted to be presupposition failures by the AtB prohibition. This supports an interpretation of the AtB prohibition as reflecting a less strict requirement, in line with a *conditional strengthening* presupposition as stated in (58).

In §5.5 we observed that the points of lowest endorsement for *unless* sentences with *most* and *few* occurred at those points where the target dot proportion precisely matched the non-target dot proportion set by the *mid* condition. If conditional strengthening requires *unless* contexts to support a salient, dot-related difference between the target and non-target marbles, we would expect this to take into account any expectation set by the quantifier—thus a *most-unless* stimulus used in a display context where the proportion of target marbles exceeded the proportion of non-target marbles would be judged worse than the same sentence in a context where the non-target marble proportion was higher, even if the target proportion was still reasonably representable as *most*. A symmetrical situation was predicted for *few*. We were unable to examine this prediction on the *mid* condition due to the pressure from 'edge' effects: for instance, target proportion 1 represents a salient difference from non-target proportion 0.8 in a *most-unless* context due to the availability of the quantifier *all* to describe the proportion of target marbles with dots. The *low* condition offers at least some opportunity to examine this prediction for *most*, as does the *high* condition for *few*. In the first case, we find the predicted pattern: endorsement rates steadily decrease for *most-unless*, non-target dot proportion 0.6, from target proportion 0.6 (20.2%) to 1 (9.8%); it is worth comparing this to results from the *high* condition, which show the same endorsement rate pattern, but where the endorsement rates for the target proportions 0.6 and 0.8—crucially, those proportions that are lower than the non-target dot proportion for this condition—are markedly higher (33.7% and 26.9%, respectively) than on their *low*-condition counterparts. The relevant *few-high* data are also compatible with the 'salient difference' prediction, though less clearly: endorsement for target proportion 0.2 is slightly decreased from endorsement in the (non-target-matching) 0.4 condition, but the 0 target proportion is in fact endorsed at a higher rate than both of these points, presumably due to the availability of the quantifier *none*, as distinct from *few*, to describe the fraction of target marbles with dots.[15]

The *mid* condition data from *some* also indicated the existence of an anti-AtB-type inference associated with *if not* conditionals; we capture this in §6.2 as the realization of the conditional strengthening (Need-a-Reason) implicature. *Some* data from both *high* and *low* conditions show the same effect: endorsement rates for *if not* group together across the target proportions 0.2 to 1, at a range of values that are lower than those for the 0 target dot proportion. Data from *some-unless* differ reliably from *some-if not* across all non-zero target dot proportions. This is broadly compatible with the predictions of presuppositional interpretation of the AtB prohibition; although there are points at which the reliable distinction does not fully seem to be categorical (e.g. target dot proportions 0.8-1 on *some-low*), these occur at points where both salient difference reasoning and the availability of an alternative quantifier for describing the target proportion are in effect.

---

15 A comparable edge effect at target proportion 1 for *most-low* is not in evidence. We do not know why this is the case, but we note that the rejection rate for *most* as a description of proportion 1 is empirically lower than the rejection rate for *few* as a description of proportion 0, and it therefore seems reasonable to expect much stronger and more marked effects from the latter case than the former.

Finally, these additional data remain compatible with the existence of biconditionality inferences, on both *if not* and *unless*. Tables 11-12 replicate the pattern in Table 8: lower average endorsement over biconditionality-challenging contexts than over biconditionality-compatible ones.

Target props.	MOST	FEW
> 0.5	60.8	**72.2**
< 0.5	**76.6**	58.8

Table 11: Average endorsement rates for *if not* in *high* non-target conditions

Target props.	MOST	FEW
> 0.5	45.7	**41.8**
< 0.5	**75.1**	31.9

Table 12: Average endorsement rates for *if not* in *low* non-target conditions

The same trend holds for *unless* statements, with larger average differences – as predicted by both the AtB condition and biconditionality—between the target proportions >0.5 and <0.5. As noted in the main text, it is difficult on the basis of the current data to fully separate the effects of a potential inference to biconditionality from the influence of conditional strengthening. We do observe, however, that the divergence of *most-unless* from *most-if not* seems to begin around target proportion 0.4, and increase as the target proportion does; when the non-target proportion is lower (see Figure 7), the divergence appears to begin earlier. In the *high* condition, divergence between *unless* and *if not* emerges between 0.6 and 0.4, and increases as the target proportion decreases. There is no reliable difference between the two conditionals in our *few-low* data, but this result remains unanalyzable due to the extremely low endorsement of *few* as a description of proportion 0 evidenced in our control data (Figure 4, Table 6).

References

Abrusán, Márta & Kriszta Szendrői. 2013. Experimenting with the King of France: Topics, verifiability, and definite descriptions. *Semantics and Pragmatics* 6. 1–43. https://doi.org/10.3765/sp.6.10.

van der Auwera, Johan. 1997. Pragmatics in the last quarter century: the case of conditional perfection. *Journal of Pragmatics* 27. 261–274.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2014. *lme4: Linear mixed-effects models using eigen and s4*. http://CRAN.R-project.org/package=lme4. R package version 1.1-7.

Chierchia, Gennaro & Sally McConnell-Ginet. 1990. *Meaning and grammar*. Cambridge, MA: MIT Press.

Clark, Herbert & Eve Clark. 1977. *Psychology and language: an introduction to psycholinguistics*. London: Thomson Learning.

Cremers, Alexandre, Manuel Križ & Emmanuel Chemla. 2016. Probability judgements of gappy sentences. Ms. LSCP & University of Vienna.

Dalrymple, Mary, Makoto Kanazawa, Sam Mchombo & Stanley Peters. 1994. What do reciprocals mean? In *Proceedings of Semantics and Linguistic Theoryemantics and linguistic theory*, vol. 4, 61–78. https://doi.org/10.3765/salt.v4i0.2466.

Dancygier, Barbara. 1985. *If, unless*, and their Polish equivalents. *Papers and Studies in Contrastive Linguistics* 20. 64–72.

Doran, Ryan, Gregory Ward, Meredith Larson, Yaron McNabb & Rachel E Baker. 2012. A novel experimental paradigm for distinguishing between what is said and what is implicated. *Language* 88. 124–154. https://doi.org/10.1353/lan.2012.0008.

Égré, Paul & Mikael Cozic. 2012. Conditionals. In Maria Aloni & Paul Dekker (eds.), *Handbook of semantics*, Cambridge: Cambridge University Press.

Fillenbaum, Samuel. 1986. The use of conditionals in inducements and deterrents. In Elizabeth Traugott, Alice ter Meulen, Judy Reilly & Charles Ferguson (eds.), *On conditionals*, 179–195. Cambridge: Cambridge University Press.

von Fintel, Kai. 1998. Quantifiers and 'if'-clauses. *Philosophical Quarterly* 48. 209–214. https://doi.org/10.1111/1467-9213.00095.

von Fintel, Kai. 1992. Exceptive conditionals: the meaning of "unless". In *Proceedings of the North East Linguistics Society*, vol. 22, 135–151.

von Fintel, Kai. 1993. Exceptive constructions. *Natural Language Semantics* 123–148.

von Fintel, Kai. 1994. *Restrictions on quantifier domains*: University of Massachusetts at Amherst dissertation.

von Fintel, Kai. 2001. Conditional strengthening: a case study in implicature. Manuscript, MIT.

von Fintel, Kai & Sabine Iatridou. 2002. If and when if-clauses can restrict quantifiers. Ms, Workshop in Philosophy and Linguistics at the University of Michigan.

Franke, Michael. 2009. *Signal to act: game theory in pragmatics*: Universiteit van Amsterdam dissertation.

Garcia-Alvarez, Ivan. 2008. *Generality and exception: a study in the semantics of exceptives*: Stanford University dissertation.

Geis, Michael. 1973. "If" and "unless". In Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli & Sol Saporta (eds.), *Issues in linguistics: papers in honor of Henry and Renee Kahane*, 231–253. Urbana, IL: University of Illinois Press.

Geis, Michael & Arnold Zwicky. 1971. On invited inferences. *Linguistic Inquiry* 2. 561–566.

Geurts, Bart. 2004. On an ambiguity in quantified conditionals. Ms, University of Nijmegen.

Grice, H. P̃aul. 1989. *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.

Higginbotham, James. 1986. Linguistic theory and Davidson's program in semantics. In Ernie Lepore (ed.), *Truth and interpretation: perspectives on the philosophy of Donald Davidson*, 29–48. Oxford: Blackwell.

Higginbotham, James. 2003. Conditionals and compositionality. *Philosophical Perspectives* 17. 181–194. https://doi.org/10.1111/j.1520-8583.2003.00008.x.

Hoeksema, Jack. 1987. The logic of exception. In Miller & Powers (eds.), *Escol1987*, Ohio: The Ohio State University.

Hoeksema, Jack. 1990. Exploring exception phrases. In Martin Stokhof & Leen Torenvliet (eds.), *Proceedings of the Seventh Amsterdam Colloquium*, 165–190.

Horn, Laurence. 2000. From if to iff: conditional perfection as pragmatic strengthening. *Journal of Pragmatics* 32. 289–326. https://doi.org/10.1016/S0378-2166(99)00053-3.

Huitink, Janneke. 2009. Quantified conditionals and compositionality. *Language and Linguistics Compass* 4. 42–53. https://doi.org/10.1111/j.1749-818X.2009.00175.x.

Janssen, Theo. 1997. Compositionality. In Johan van Benthem & Alice ter Meulen (eds.), *Handbook of logic and linguistics*, 417–475. Amsterdam: Elsevier.

Keenan, Ed & Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9. 253–326.

Kratzer, Angelika. 1986. Conditionals. In *Proceedings of the Chicago Linguistic Society*, vol. 22, 1–15.

Kratzer, Angelika. in press. Chasing Hook: Quantified indicative conditionals. In Lee Walters & John Hawthorne (eds.), *Conditionals, probability, and paradox: Themes from the philosophy of dorothy edgington*, .

Lauer, Sven. 2013. *Towards a dynamic pragmatics*: Stanford University dissertation.

Leslie, Sarah-Jane. 2009. "If," "unless," and quantification. In Robert Stainton & Christopher Viger (eds.), *Compositionality, context, and semantics*, vol. 85 Studies in Linguistics and Philosophy, 3–30. Dordrecht: Springer.

Levinson, Stephen. 2000. *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge MA: MIT Press.

Levy, R. 2014. Using R formulae to test for main effects in the presence of higher-order interactions. *ArXiv e-prints* http://adsabs.harvard.edu/abs/2014arXiv1405.2094L.

Löbner, Sebastian. 2000. Polarity in natural language: predication, quantification and negation in particular and characterizing sentences. *Linguistics and Philosophy* 23. 213–308.

Malamud, Sophia. 2012. The meaning of plural definites: a decision-theoretic approach. *Semantics and Pragmatics* 5. 1–58. https://doi.org/10.3765/sp.5.3.

Moltmann, Friederieke. 1995. Exception sentences and polyadic quantification. *Linguistics and Philosophy* 18. 223–280.

Moxey, Linda & Anthony Sanford. 1993. *Communicating quantities: a psychological perspective.* Mahwah, NJ: Laurence Erlbaum Associates.

Nadathur, Prerna. 2013. If...(and only if): conditional perfection and completeness. University of Oxford, MPhil thesis.

Nadathur, Prerna. 2014. *Unless*, exceptionality, and conditional strengthening. In Ronald de Haan (ed.), *Proceedings of the ESSLLI 2014 Student Session*, 109–120.

Nadathur, Prerna & Daniel Lassiter. 2014. Unless: an experimental approach. In *Proceedings of Sinn und Bedeutung*, vol. 19, 426–444.

Nadathur, Prerna & Daniel Lassiter. 2018. An experimental look at the negative implications of exceptives. In *Proceedings of the North East Linguistics Society*, vol. 48, 201–210.

Nouwen, Rick. 2003. Complement anaphora and interpretation. *Journal of Semantics* 20. 73–113. https://doi.org/10.1093/jos/20.1.73.

Peters, Stanley & Dag Westerståhl. 2008. *Quantifiers in language and logic,* chap. 8: Exceptive quantifiers, 297–323. Oxford: Oxford University Press.

Potts, Christopher. 2005. *The logic of conventional implicatures.* Oxford: Oxford University Press.

Prochnow, Herbert. 1955. *Speaker's handbook of epigrams and witticisms,* Speakers' & Toastmasters' Library. New York: Harper.

Quine, Willard. 1959. *Methods of logic*. New York: Holt, Rinehart & Winston.

R Core Team. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. http://www.R-project.org/.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: The Free Press.

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

Schwarz, Florian. 2016. False but slow: Evaluating statements with non-referring definites. *Journal of Semantics* 33. 177–214. https://doi.org/10.1093/jos/ffu019.

Schwarzschild, Roger. 1991. *On the meaning of definite plural noun phrases*: University of Massachusetts–Amherst dissertation.

Szabó, Zoltan. 2008. Compositionality. In *The Stanford Encyclopedia of Philosophy*, Winter 2008 edition.

Szabolcsi, Anna & Bill Haddican. 2004. Conjunction meets negation: a study of cross-linguistic variation. *Journal of Semantics* 21. 219–250. https://doi.org/10.1093/jos/21.3.219.

Zaefferer, Dietmar. 1991. Conditional and unconditionals: cross-linguistic and logical aspects. In Dietmar Zaefferer (ed.), *Semantic universals and universal semantics*, Berlin: Foris Publications.

Prerna Nadathur
Department of Linguistics, Stanford University
pnadathur@stanford.edu

Daniel Lassiter
Department of Linguistics, Stanford University
danlassiter@stanford.edu