

INDEPENDENCE DAY?

1 INTRODUCTION

Two recent and influential papers, van Rooij 2007 and Lassiter 2012, propose solutions to the proviso problem that make central use of related notions of independence—qualitative in the first case, probabilistic in the second. We argue here that, if these solutions are to work, they must incorporate an implicit assumption about presupposition accommodation, namely that accommodation does not interfere with existing qualitative or probabilistic independencies. We show, however, that this assumption is implausible, as updating beliefs with conditional information does not in general preserve independencies. We conclude that the approach taken by van Rooij and Lassiter does not succeed in resolving the proviso problem.

2 THE PROVISIO PROBLEM

Standard theories of semantic presupposition,¹ in particular satisfaction theory, predict that the strongest semantic presupposition of an indicative conditional with the form of (1) (where B_p is a sentence with a presupposition P) is the material conditional (2):²

- (1) If A then B_p
 (2) $A \supset P$

These theories make similar predictions for presupposition triggers embedded in right conjuncts and disjuncts (with ‘ \rightsquigarrow ’ meaning ‘presupposes’):

Mandelkern: All Souls College, Oxford, OX1 4AL

Rothschild: University College London, Gower Street, London, WC1E 6BT

The authors contributed equally. Many thanks to Daniel Lassiter, Rick Nouwen, Benjamin Spector, and two anonymous referees for this journal for very helpful comments.

¹With the exceptions of approaches in DRT [van der Sandt, 1989, 1992, Geurts, 1996, 1999, Kamp, 2001] and dissatisfaction theory [Mandelkern, 2016a], but including satisfaction theory (our focus), spelled out in Heim 1982, 1983, 1990, 1992, based on earlier work in Stalnaker 1973, 1974, Karttunen 1974, and since developed in Beaver 2001, von Stechow 2008, Schlenker 2009, Lassiter 2012, Rothschild 2011/2015, and others, as well as multivalent theories—see Kleene 1952, Strawson 1952, van Fraassen 1969, Peters 1977, Karttunen and Peters 1979, and more recently George 2008, Fox 2012—and a heterogeneous variety of other theories, e.g., Soames 1982, Schlenker 2008, Chemla 2008.

²Capital roman letters stand for sentences; corresponding italic letters stand for the propositions they express (suppressing reference to contexts for readability). Following van Rooij and Lassiter, we focus on indicative conditionals here, though similar problems arise for subjunctives.

- (3) A and B_p
 $\rightsquigarrow A \supset P$
- (4) A or B_p
 $\rightsquigarrow \neg A \supset P$

Are these predictions correct? According to satisfaction theory, the presuppositions of a sentence are contents that an input context must entail before the sentence can be added to the context. Consider an example (modified slightly) from Geurts 1996:

- (5) If Theo hates sonnets, then so does his wife.

According to the predictions just reviewed, given that ‘his [Theo’s] wife’ presupposes that Theo has a wife, (5) presupposes the material conditional (6) (perhaps most naturally realized in natural language as the truth-conditionally equivalent disjunction in (7)):

- (6) Theo hates sonnets \supset Theo has a wife.
 (7) Theo doesn’t hate sonnets or Theo has a wife.

Now consider a context which entails (7). One way to imagine such a context is to imagine, for instance, that Mark has asserted (7) and it has been accepted in the context. Then suppose that Susie asserts (5). According to satisfaction theory, the presupposition of (5) is satisfied in this context, and so we will not be required to *accommodate* any new presuppositions when (5) is asserted—that is, we will not need to quietly adjust the input context in order to make sure that (5)’s presuppositions are all satisfied. And indeed, this seems to be just what we observe. In particular, in such a context, an assertion of (5) is not felt to presuppose that Theo has a wife. One classic test for presupposition is the ‘Hey wait a minute’-test [von Stechow, 2008]: in the context in question, responding to (5) with something like ‘Hey wait a minute! I didn’t know Theo had a wife!’ feels like a non sequitur. This result is consistent with the predictions of satisfaction theory.

But problems arise when we consider an assertion of (5) in a context that *does not* already entail its presupposition in (6). For instance, consider a context in which nothing is known about Theo, and suppose that Susie asserts (5). Susie will ordinarily be felt to be presupposing not just the material conditional (6), but also its unconditional consequent: that Theo has a wife. The easiest way to see this, again, is to note that, in *this* context—unlike the one just imagined—a response to (5) with ‘Hey wait a minute! I didn’t know Theo has a wife!’ feels entirely appropriate. This suggests that the presupposition we accommodate in this context—what we add to the common ground to render Susie’s assertion felicitous—is something stronger than the predicted conditional presupposition

(6) of (5), and instead is the unconditional consequent of (6).

Examples like this one have led to a consensus in the literature that the predictions of satisfaction theory about the presupposition of (5) seem perfectly adequate when we focus on *contexts which already entail that presupposition*. But when we look at contexts in which something must be accommodated, it looks as though we are inclined to accommodate something much stronger than what satisfaction theory predicts we have to accommodate. Similar issues arise for conjunctions and disjunctions.

This gap between the predicted *conditional* semantic presuppositions of complex sentences embedding presupposition triggers, on the one hand, and the observed *unconditional* propositions which we accommodate when the conditional presupposition is not already common ground, on the other, is the proviso problem. The name was given by Geurts [1996], but the problem has been recognized at least since Karttunen and Peters 1979. Let us emphasize that this is a problem about *accommodation* in particular. Geurts [1996, p. 270–271] is particularly clear about the link between the proviso problem and accommodation: he argues that *if* the satisfaction theory could only make the correct prediction about *accommodation* in cases like (5) then there would be no proviso problem. That is, the predictions made by satisfaction theories which are apparently problematic are *just* its predictions about accommodations. Heim 2006 puts this point particularly clearly: ‘When the predicted conditional presupposition is in the common ground, the [relevant] sentences are felicitous and don’t require additional accommodation. The [problematic] judgments. . . are judgments about what we spontaneously accommodate when presented with out-of-the-blue utterances.’ (See also von Stechow 2008, 160–161 for the same point.) We emphasize this here because neither van Rooij or Lassiter frames the proviso problem as a problem about accommodation in particular.³ In this paper we adapt their views slightly—and charitably, we hope—to explicitly cover the case of accommodation (i.e. to cover cases in which the context before the utterance is not one in which the speaker can assume the listener has already accepted the presupposition).⁴

Before proceeding, we should note that, in two important classes of cases, the predictions of satisfaction theory seem to be vindicated. The first class comprises sentences like (8):

(8) If Theo has a wife, his wife hates sonnets.

Satisfaction theory predicts that the presupposition of (8) is ‘Theo has a wife \supset Theo has a wife’, which is, of course, a tautology, and thus satisfaction theory predicts that (8) has no non-trivial presuppositions. This prediction seems to be

³Thanks to two anonymous referees for noting this point.

⁴In Lassiter’s framework, these are cases in which the speaker cannot assume the listener has already conformed their credences to the presuppositional requirements.

correct: there are no contexts in which someone who asserts (8) will be felt to presuppose that Theo has a wife. (This point, while important to keep in mind, will not play much of a role in what follows.)

Second, in *some* cases, we *do* seem to accommodate the predicted conditional presupposition of satisfaction theory. Consider (9):⁵

(9) If Buganda is a monarchy, then Buganda’s king will be at the meeting.

If someone asserts (9) in a null context, we will typically accommodate (10), *not* (11):

(10) Buganda is a monarchy \supset Buganda has a king.

(11) Buganda has a king.

(10) is non-trivial—monarchies can have queens as well—and yet in this case we really do seem to accommodate a conditional, whereas in (5) we did not. An adequate solution to the proviso problem must make sense not only of the fact that, in cases like (5), we accommodate something stronger than the predicted conditional presupposition, but also of the fact that, in cases like this one, we accommodate just the conditional.

3 TWO NOTIONS OF INDEPENDENCE

Van Rooij [2007] and Lassiter [2012] propose closely related solutions to the proviso problem: accounts which aim to explain why we accommodate something unconditional in cases like (5) (if we have to accommodate something at all), but only a conditional in cases like (9). Both accounts make crucial use of a notion of independence. We begin our discussion of these accounts by defining the two notions and discussing the relationship between them.

We begin with van Rooij’s notion of QUALITATIVE INDEPENDENCE (our term). Van Rooij’s paper in fact defines two distinct notions of independence. Van Rooij claims these are equivalent, but, as we show in Appendix A, they turn out not to be. However, as far as we can tell, this fact is inessential to the main thrust of van Rooij’s theory, and so we think the most charitable option is to ignore the first notion of independence and focus on the second; see Appendix A for discussion of why we think this is the right notion of the two to focus on. Apart from this (we hope charitable) emendation—and the generalization of the account to cover accommodation explicitly—our summary of van Rooij’s theory is intended to be faithful to his presentation.

Van Rooij is working in a standard Boolean framework in which a background context can be characterized as a set of possible worlds from a stock of worlds W , and propositions are also subsets of W . Van Rooij adopts satisfaction theory as his theory of semantic presupposition, on which presuppositions are propositions

⁵Following many similar examples in the literature; see Geurts 1996, Beaver 2001.

which must be entailed by their input contexts. The central notion in van Rooij's account is that of the qualitative independence of A and B , relative to a context s .⁶

Propositions A and B are QUALITATIVELY INDEPENDENT in a context s iff

- a) if $A \cap s \neq \emptyset \wedge B \cap s \neq \emptyset$ then $A \cap B \cap s \neq \emptyset$
- b) if $A^c \cap s \neq \emptyset \wedge B \cap s \neq \emptyset$ then $A^c \cap B \cap s \neq \emptyset$
- c) if $A \cap s \neq \emptyset \wedge B^c \cap s \neq \emptyset$ then $A \cap B^c \cap s \neq \emptyset$
- d) if $A^c \cap s \neq \emptyset \wedge B^c \cap s \neq \emptyset$ then $A^c \cap B^c \cap s \neq \emptyset$.

The notion of independence that plays a key role for Lassiter is closely related, but is probabilistic rather than qualitative. Lassiter argues that the operative notion of an information state in the theory of presupposition should not be a context (a set of possible worlds), but rather a probability function over a set of possible worlds. Lassiter proposes that presuppositions need not be entailed by their input context, in the standard qualitative sense, but rather must have high probability (higher than a threshold t) in their local information state. The notion of independence that plays a key role in Lassiter's theory is just the standard probabilistic one:

Propositions A and B are PROBABILISTICALLY INDEPENDENT relative to a probability function p iff $p(A \cap B) = p(B)p(A)$.

There is a precise sense in which Lassiter and van Rooij's notions are related: for a finite set of worlds W , subsets A and B of W are qualitatively independent in a context s just in case there is a probability space built on W which assigns non-zero probability to all and only the worlds of s , and which makes A and B probabilistically independent. The proof is in Appendix B. At a high level, this fact shows that there is a sense in which qualitative independence is the corollary of probabilistic independence in a qualitative framework, though also a sense in which it is weaker.

4 FROM INDEPENDENCE TO STRENGTHENING

Now we are in a position to see how van Rooij and Lassiter use their notions of independence to try to solve the proviso problem.

In both cases we apply their accounts to cases in which a presupposition needs to be accommodated rather than cases in which a presupposition is already accepted in the context. We do this because, as we argued in Section 2, these cases which are problematic for a satisfaction-style theory. It is in these cases that the satisfaction theory struggles to explain why we accommodate unconditional presuppositions when conditional presuppositions should be sufficient.

⁶Where P is a proposition, P^c is its complement in W , i.e. $W \setminus P$.

We begin with van Rooij, who, recall, is working in the framework of satisfaction theory. Suppose the conditional ‘If A , then B_p ’ is asserted in a context c which does not already include $A \supset P$. This, again, is the case that is of interest to us, since there is no problem for satisfaction theory in cases in which $A \supset P$ is already entailed by the input context. According to satisfaction theory, this presupposition will need to be accommodated: quietly added to the context so that we can process the assertion. In particular, the context after accommodation (call it c') will have to entail $A \supset P$. Van Rooij assumes, following Stalnaker 1975, that a conditional can only be asserted when its antecedent is compatible with the context, so A is compatible with c ; van Rooij is not explicit about it, but he must be assuming that A must remain compatible with c through the process of accommodation, so A is also compatible with c' . Now suppose that A and P are qualitatively independent in c , and *remain independent* in the posterior context c' . The only way these conditions can all be met is if c' entails P . Otherwise, there would be P^c worlds in c' , and since we know there are A worlds in c' , there would have to be $A \cap P^c$ worlds in c' , by the qualitative independence of A and P , contrary to the assumption that c' entails $A \supset P$.

In short, if A and P are qualitatively independent in a context, *and that qualitative independence is preserved through the process of accommodation* when ‘If A , then B_p ’ is asserted, then P , and not just $A \supset P$, will be entailed by the posterior context.

And how does this story help with the proviso problem? Van Rooij’s idea is that, as a matter of empirical fact, the conditional presuppositions which we tend to strengthen are just those which are qualitatively independent in most contexts. To compare the two key examples from above, *Theo hates sonnets* will typically be treated as qualitatively independent of *Theo has a wife*, and so we will generally strengthen ‘*Theo hates sonnets* \supset *Theo has a wife*’ to ‘*Theo has a wife*’ when we have to accommodate this presupposition. By contrast, *Buganda has a king* will often *not* be treated as qualitatively independent of *Buganda has a monarchy*, since there will be no worlds where Buganda fails to be a monarchy but still has a king; and so ‘*Buganda is a monarchy* \supset *Buganda has a king*’ is correctly predicted to remain unstrengthened.

Lassiter’s story is similar to van Rooij’s, albeit in a slightly different background setting (and with some interesting empirical differences on which, however, we will not focus).⁷ Lassiter’s theory of presupposition is a probabilistic variant of satisfaction theory: when ‘If A , then B_p ’ is asserted, on Lassiter’s theory, speakers must ensure that the conditional probability of P on A is suitably high— that is, is above some threshold t .⁸ In this framework, the proviso problem arises

⁷A similar idea to Lassiter’s can be found in Singh 2007, 2008, Schlenker 2011.

⁸Lassiter thinks speakers can only assert something if they believe its presuppositions are given a high conditional probability by themselves as well as by their audience (as in his (9) on p. 10). So we can, loosely, think of Lassiter as also putting a condition on the ‘context’ or ‘common ground’. We speak in this way throughout the paper in order to emphasize the connections between Lassiter’s paper and traditional

in essentially the same form as for satisfaction theory: in many cases in which we must change the input context to meet this presuppositional requirement, it seems as though we do so by not only setting the conditional probability of P on A high, but also setting the *unconditional* probability of P high. This will be so in the case of Geurts' (5) (we not only set the conditional probability of *Theo has a wife* on *Theo hates sonnets* high; we also set the unconditional probability of *Theo has a wife* high). In other cases, however, as in the case (9), it seems that we only update by setting the conditional probability of P on A high, without making P unconditionally probable (we set the probability of *Buganda has a king* on *Buganda is a monarchy* high, without making it unconditionally probable that Buganda has a king).

Lassiter proposes to make sense of this situation as follows. Suppose that \ulcorner If A , then $B_p \urcorner$ is asserted in a context in which the conditional probability of P on A is not yet high. On Lassiter's theory, speakers must change the context—they must 'accommodate,' to extend the ordinary usage of that word—to ensure that the conditional probability of P on A is above a threshold t . It follows that, relative to the posterior probability function p' (the probability function that results from this accommodation), $p'(P|A) \geq t$. Now suppose further that A and P are probabilistically independent in c , and that *this probabilistic independence is preserved through the process of accommodation*. Now, by the assumption that P and A are probabilistically independent under p' , and the standard definition of conditional probability⁹ (assuming $p(A) > 0$), $p'(P|A) = \frac{p'(P \cap A)}{p'(A)} = \frac{p'(P)p'(A)}{p'(A)} = p'(P)$. That is, when P and A are probabilistically independent under p' , the conditional probability of P on A under p' is equal to the unconditional probability of P under p' . So, since $p'(P|A) \geq t$, it follows that $p'(P) \geq t$. Thus the posterior context will support P (in the probabilistic sense of assigning suitably high probability), and not just $A \supset P$. In short, if A and P are probabilistically independent in a context, and that probabilistic independence is preserved through the process of accommodation, then the posterior context will support P , and not just $A \supset P$.

This solution promises to make sense of the contrast between our two key examples in very similar ways to van Rooij's proposal. In most contexts, *Theo hates sonnets* will be treated as probabilistically independent of *Theo has a wife*; so when we update the context to make the conditional probability of the latter on the former high, we will also thereby update to make the *unconditional* probability of *Theo has a wife* high. By contrast, *Buganda has a king* and *Buganda is a monarchy* will *not* in most contexts be treated as probabilistically independent: the latter will usually probabilify the former. So when we update contexts to make the conditional probability of the former on the latter high, we will typically *not* strengthen to make the unconditional probability of the former on the latter

satisfaction theory, though the common ground here need not have the iterated structure characteristic of the common ground in standard formulations of satisfaction theory—an issue which does not bear on our claims here.

⁹On which $p(A|B) = \frac{p(A \cap B)}{p(B)}$ provided $p(B) > 0$.

high.

So far, this account is very similar to van Rooij's, albeit with a probabilistic rather than qualitative notion of independence. In some other cases, van Rooij's and Lassiter's accounts diverge. One point of departure comes from cases where qualitative and probabilistic independence come apart. A second divergence comes from cases in which A and P are not intuitively independent in any sense, but rather A *decreases* the probability of P —that is, $p(P|A) < p(P)$. In that case, assuming, again, that this property of probability distributions is preserved across accommodation, if the posterior probability function p' has $p'(P|A) \geq t$, then, since $p'(P) \geq p'(P|A)$, we will also have $p'(P) \geq t$. Lassiter gives examples that suggest that, indeed, we strengthen presuppositions not only when the antecedent and consequent are independent, but also when the antecedent disprobabilizes the consequent.¹⁰ For this reason, Lassiter takes his account to be more general than independence-based accounts such as van Rooij's.

5 THE PROBLEM: PRESERVATION THROUGH ACCOMMODATION

These stories are elegant and insightful, and there is much to like about them. But our exposition makes clear that both approaches make a crucial assumption: that the relevant kinds of independence (and, in Lassiter's case, disprobabilization) properties will generally be *preserved across presupposition accommodation*.^{11,12} Thus, for instance, focusing first on van Rooij's theory, we want to predict that when we accommodate the conditional presupposition (6), it will be strengthened to the unconditional (12):

(6) Theo hates sonnets \supset Theo has a wife.

(12) Theo has a wife.

To predict this strengthening, we must assume that *Theo has a wife* is qualitatively independent of *Theo hates sonnets*. This assumption perfectly plausible for most 'default' contexts. In general, there is no reason to think that, in a given context, we would, for example, leave open that Theo does or does not hate sonnets, but assume that either he doesn't hate sonnets, or he has a wife (one way to violate qualitative independence). And, likewise, in general, there is no reason to think Theo hating sonnets is probabilistically relevant to his having a wife.

But *these* assumptions are not yet sufficient to account for the strengthening of (6) to (12). What we need, moreover, is the claim that this independence assumption persists *even after* the context is updated with the material conditional (6) by way of presupposition accommodation. Without this assumption, again,

¹⁰ P disprobabilizes Q relative to p iff $p(Q|P) < p(Q)$.

¹¹We here again note that this assumption arises only as part of our extension of their accounts to deal with the crucial case of accommodation; again, they themselves do not explicitly discuss such cases, and hence do not directly address the proviso problem.

¹²See Franke 2007, Francez 2015, Goebel 2017 for related points about different formulations of independence conditions.

we do not have what we need. The qualitative independence of *Theo hates sonnets* and *Theo has a wife* in the *antecedent* context does not yet get us strengthening; it is only if these remain independent after (6) is accommodated that we know (6) will be strengthened to (12).

Likewise, in Lassiter's framework, we want to predict that when we adjust the input context to make the conditional probability of *Theo has a wife* on *Theo hates sonnets* high, we will *also* make the unconditional probability of *Theo has a wife* high. To predict this, we must assume that *Theo has a wife* is probabilistically independent of (or disprobabilized by) *Theo hates sonnets*. As an assumption about most 'default' contexts, the former of these assumptions—that these propositions are treated as probabilistically independent—seems perfectly plausible. But *that* assumption, again, does not yet suffice to ensure that, when we accommodate to make the conditional probability of *Theo has a wife* on *Theo hates sonnets* high, we also make the unconditional probability of *Theo has a wife* high. What we need is the additional assumption that these propositions *remain* probabilistically independent after this update. It is only if we make this assumption that we can conclude that the posterior context assigns high probability to *Theo has a wife*.

But what reason do we have to think these will remain independent after accommodation? One possibility is that van Rooij and Lassiter are implicitly making the following closely related assumptions, respectively:

QUALITATIVE RESPECT: If A and P are qualitatively independent in a given context and we accommodate a presupposition of the form $A \supset P$, A and P will remain qualitatively independent in the posterior context.

PROBABILISTIC RESPECT: If A and P are probabilistically independent (resp. A disprobabilizes P) in an input context, and we accommodate a presupposition that the probability of P on A is high, then A and P will remain probabilistically independent (resp. A will still disprobabilize P) in the posterior context.

If these RESPECT principles could be justified, then this would explain why, in van Rooij's framework, when A and P are qualitatively independent, we generally strengthen $A \supset P$ to P when we accommodate it; and, in Lassiter's framework, when A and P are probabilistically independent, or A disprobabilizes P , we generally make the probability of P high when we accommodate a presupposition that the conditional probability of P on A is high. And, given the discussion so far, this result is just what we need to solve the proviso problem (at least *modulo* the further issues discussed in the conclusion).

But what could justify these RESPECT principles? Van Rooij and Lassiter do not discuss them explicitly, and it is hard for us to find a conceptually respectable

foundation for either one.

The first point to make here is that presupposition accommodation is just supposed to be change in beliefs (qualitative or probabilistic), and changes in beliefs do not *in general* respect independences. To take a simple case, suppose I think that *Bill goes to the party* and *Sue goes to the party* are qualitatively and probabilistically independent, and that I have no idea whether either is true. Suppose further I don't know whether Bill and Sue are dating. Then I learn they are dating and go everywhere together. Then I should no longer take *Bill goes to the party* and *Sue goes to the party* to be either qualitatively or probabilistically independent. I should not leave it open that one goes to the party while the other does not. And I should judge the probability that Sue goes, conditional on the probability that Bill goes, to be much higher than the unconditional probability that Bill goes.

Take another case. Suppose we don't know if either Bill or Ted is going to the party, and assume that each going is probabilistically independent of the other going. We are then told simply that the conditional probability is high that Bill goes, conditional on Ted going. We thus raise our conditional probability of Bill going, conditional on Ted going. It would be downright weird in this case to assume probabilistic independence is preserved—something which would force us to have high probability in the unconditional proposition that Bill goes. Nothing in what we have learned would justify that.

In addition, updates do not generally maintain disprobabilization. Suppose I think that Bill and Ted dislike each other, so that *Bill goes to the party* makes *Ted goes to the party* less likely, and I antecedently have no idea whether either of them will go the party. Now I learn that Ted is in fact certain to go to the party if Bill does. There is nothing wrong with this update. But in my posterior probability state, *Bill goes to the party* obviously no longer disprobabilizes *Ted goes to the party*.

It thus is not true that updates preserve independence properties, or disprobabilization properties, *in general*. And, worse, it is precisely updates with material conditionals or conditional probabilities which seem to be prime candidates for disrupting independence properties in particular. To return to our example from the beginning, suppose that *Theo hates sonnets* and *Theo has a wife* are antecedently qualitatively independent, and you don't know whether either is true. Now suppose you learn the material conditional *Either Theo doesn't hate sonnets, or he has a wife*. Should the two disjuncts remain qualitatively independent? Intuitively not. You still don't know whether Theo hates sonnets and whether he has a wife. But you can now rule out worlds where he both hates sonnets and doesn't have a wife. That means that the two propositions are no longer qualitatively independent.

Likewise, suppose that *Theo has a wife* and *Theo hates sonnets* are antecedently probabilistically independent. Then you learn that the conditional probability

of *Theo has a wife* on *Theo hates sonnets* is high (perhaps by way of learning a conditional like *If Theo hates sonnets, then he is very likely to have a wife*). You now have high credence in Theo having a wife, conditional on his hating sonnets. But should you preserve probabilistic independence, so that you *also* have high credence that Theo has a wife? There is clearly no reason for you to do so: intuitively, you haven't learned anything about the chances that Theo has a wife.

We can make these considerations more precise and general as follows. First consider the qualitative case. Suppose A and B are qualitatively independent, with all of A , A^c , B , and B^c compatible with your information state. Then you learn $A \supset B$ by adding that content in a *minimal* way to your antecedent attitude state. In a standard qualitative framework for modeling belief revision (like AGM theory [C.E. Alchourrón and P. Gärdenfors and D. Makinson, 1985]), the minimal update amounts to removing just the $A \cap B^c$ worlds from your information state. Such an update is *guaranteed* to disrupt the qualitative independence of A and B : they will no longer be qualitatively independent on the posterior context, since there will still be A and B^c worlds, but no $A \cap B^c$ worlds.

The overall picture is the same in the probabilistic case, though the details are more complex. Lassiter, again, takes contexts to be probability functions rather than sets of worlds, and, again, a conditional with the form \ulcorner If A , then $B_p \urcorner$ will have a presuppositional constraint of the form $p(P|A) \geq t$, where p is the probability function of the context and t is a high threshold. There is no consensus about how to update an arbitrary probability function to meet a new condition like this, for, importantly, this condition is non-propositional: it does not amount to learning a proposition (that is, event), and thus does not amount merely to just conditioning on that proposition.¹³ However, given the intuitions elicited above, any plausible rule we adopt to cover this kind of update will not in general maintain the probabilistic independence of A and P .

Let us briefly consider two prominent such rules for updating one's probability function to respect new information about conditional probabilities. One is to choose a new probability function that minimizes the relative entropy (the *Kullback-Leibler divergence*) from the original function (see Kullback and Leibler 1951; this rule is sometimes called *infomin*, following van Fraassen 1981). The second rule is suggested by Douven and Romeijn [2011] (following a proposal in Bradley 2005) in response to van Fraassen's [1981] Judy Benjamin problem, and is closely related: this rule minimizes not relative entropy, but rather a closely related quantity, namely inverse relative entropy. Both rules are laid out fully in

¹³Jeffrey conditionalization is also not applicable here since one must update a conditional probability rather than a simple probability. Lassiter himself briefly makes use of graphical models in representing presuppositional updates, and it is natural to look to graphical models for guidance on this question. But, first, Lassiter makes explicit that the graphical models are just an expositional tool, and are insufficiently expressive to model probabilistic presupposition update. Second, graphical models are simply representations of probability distributions: they do not come with a special update rule. Thus, the problem of how to update graphical models is just the general problem of how to update probability distributions.

Appendix C.

The key fact about both rules for present purposes is that both require that, when we update our probability function to change the probability of P given A , we ought not to change the conditional probability of P given A^c (provided the probability of A remains non-maximal, that is, not 0 or 1). This constraint is very plausible on independent intuitive grounds. Suppose you start out with equal credence in the propositions that Theo hates sonnets and has a wife, that he hates sonnets and doesn't have a wife, that he likes sonnets and has a wife, and that he likes sonnets and doesn't have a wife. You learn just that the probability that Theo has a wife is high conditional on Theo hating sonnets. What should you think about the probability that he has a wife conditional on Theo *not* hating sonnets? Intuitively, this conditional probability should not change at all; you have learned nothing about the chances that he has a wife conditional on him not hating sonnets.¹⁴ This intuition generalizes: learning just something about the numerical range of the conditional probability of P on A does not generally tell us anything about the conditional probability of P on A^c . But any update rule which satisfies this constraint—which ensures that when we change the conditional probability of P on A , we leave the conditional probability of P on A^c unchanged—ensures that, if we start with a probability state in which A and P are probabilistically independent and the probability of A is non-maximal, and then update to raise the conditional probability of P on A , we will always get a new state in which A and P are *not* independent. Updating conditional probabilities in a way that respects this very intuitive constraint is thus *guaranteed* to disrupt probabilistic independences. More detailed discussion of both rules, and a proof of this fact, are found in Appendix C.¹⁵

Thus, coming back to the proviso problem, not only is it the case that updates *in general* do not preserve qualitative or probabilistic independence; worse, updates with material conditionals and updated conditional probabilities are prime candidates for disrupting independence properties.

Having said that, it is of course still *possible* to update with a material conditional or update conditional probabilities in such a way that we maintain independence (and disprobabilization) properties, by updating in a non-standard (and in some obvious sense non-minimal) way. And, if the RESPECT principles are to be defended, we will have to claim that we do just this when it comes to presupposition accommodation. That is, we will have to claim that, although updating with material conditionals and updating conditional probabilities does not *generally* respect independence, it does so when we are *accommodating* a material conditional/high conditional probability. In those cases, the idea would

¹⁴At least if we follow the intuitions put forth by van Fraassen [1981] and Bradley [2005]. What we actually do in any case will depend on the fine details of the case: this formal rule is a kind of posited default.

¹⁵We will not discuss when updates to conditional probabilities preserve disprobabilization properties, because the failure to preserve independence properties suffices for our point here.

be, we choose a non-standard update in order to ensure that we preserve the relevant independence properties.

But why would this be so? We don't see any independently motivated reason to adopt this hypothesis, or any respectable way to build it into a semantic and pragmatic system. Nor do we see any way in which this dialectical move makes progress on the proviso problem. The proviso problem, again, is the problem of accounting for stronger-than-expected updates when it comes to accommodating conditional presuppositions/high conditional probabilities. And the problem of justifying the RESPECT principles is also the problem of accounting for stronger-than-expected updates when it comes to accommodating conditional presuppositions/high conditional probabilities. But this is just the proviso problem again!

In short: Without the RESPECT principles, the views under consideration have no empirical plausibility: they fail to make sense of the core cases that they are designed to capture. The empirical ambitions of the views can be vindicated if we assume the RESPECT principles. But as a theoretical matter, defending the RESPECT principles seems very difficult—indeed, the problem of defending these principles seems to just be the proviso problem.

6 CONCLUSION

We thus do not think that van Rooij and Lassiter's proposed responses to the proviso problem are successful as they stand: these proposals contain a serious lacuna which we do not see a ready way to bridge.

In concluding, let us note a limit to the scope of our criticism. Schlenker [2011], following Singh [2006], helpfully divides the proviso problem into two (potentially separate) problems:

- (i) *Strengthening Problem*: By which mechanism can conditional presuppositions be strengthened?
- (ii) *Selection Problem*: How does one choose among the unstrengthened and strengthened presuppositions?

From the point of view of this taxonomy—which is certainly not forced on us (and which Lassiter and van Rooij do not themselves adopt), but may be helpful for situating our criticism—both van Rooij [2007] and Lassiter [2012] attempt to address both problems: as we have seen, they *select* presuppositions to be strengthened on the basis of independence (plus, in Lassiter's case, disprobabilization), and *strengthen* them using an accommodation mechanism that (they implicitly assume) satisfies the RESPECT principles. Here we have only shown that their answer to the *Strengthening Problem* is inadequate as it stands. For all we have said, it may well be that independence (of one form or another) still plays a crucial role in the selection problem: determining which presuppositions to strengthen. One place for independence to play a crucial role may be in a theory

like that advanced by Beaver [1992, 2001], on which listeners have plausibility orderings of some kind over different possible *contexts*, and use those orderings to decide which context to update to. Independence-based considerations may well play a role in that plausibility ordering. This role will be more indirect than the role played by independence-based considerations in van Rooij’s and Lassiter’s approach, and our criticism of van Rooij’s and Lassiter’s approach does not touch Beaver’s approach, which we will not try to evaluate here. Having said that, there are, however, also significant challenges already in the literature to using independence for the selection problem. One comes from Geurts [1996], who points out that even apparently independent material conditionals are not strengthened when they are presupposed in the scope of a factive attitude verb. Another comes from Gazdar [1979], Geurts [1996], Mandelkern [2016b], who note that conditional presuppositions are often strengthened even when they are intuitively *not* independent, and that this strengthening cannot be easily cancelled, casting doubt on the basic idea that the strengthening is pragmatic in nature. There is much to say about all of these challenges; we raise them here mainly to distinguish them from our own and review what we take to be the state of play for independence-based approaches to the proviso problem.

In short: while there may well be a role for independence in a solution to the proviso problem, there is substantially more work to be done to show that it can play the kind of role that van Rooij and Lassiter envision for it.

A VAN ROOIJ’S (2007) NOTION OF INDEPENDENCE

Here we show that the two different ways in which van Rooij [2007] spells out the notion of independence turn out to be inequivalent, *pace* van Rooij, and briefly justify our decision to focus on the second notion. van Rooij 2007 first defines *orthogonality of questions* (we have changed some terminology and notation to bring this in line with our own):

ORTHOGONALITY OF QUESTIONS:

Let Q_1^P and Q_2^P be two partitions, then we say that Q_1^P and Q_2^P are orthogonal with respect to each other iff $\forall q_1 \in Q_1^P : \forall q_2 \in Q_2^P : q_1 \cap q_2 \neq \emptyset$.

Van Rooij then defines the *question whether A*, and then defines the notion of *question independence*:

QUESTION WHETHER A:

The question whether A in context s (denoted $A?_s$) is the partition $\{A \cap s, A^c \cap s\}$.

QUESTION INDEPENDENCE OF A AND B IN CONTEXT s :

Formulae A and B are question independent of each other in context s iff $A?_s$ and $B?_s$ are orthogonal to each other.¹⁶

Van Rooij then defines the notion of *qualitative independence*, which is the notion we define in the main text. He then claims the following:

INDEPENDENCE LEMMA

Formula A and B are question independent of each other in context s iff they are qualitatively independent of each other in s .

But INDEPENDENCE LEMMA is false: the notion of question independence and qualitative independence come apart. To see this, consider a non-empty context s which entails A and entails B . Then (1) $A \cap s \neq \emptyset$ and $B \cap s \neq \emptyset$ are both true, and so is $A \cap B \cap s \neq \emptyset$; and (2) $A \cap s \neq \emptyset$ and $B^c \cap s \neq \emptyset$ are not both true; and (3) $A^c \cap s \neq \emptyset$ and $B \cap s \neq \emptyset$ are not both true; and (4) $A^c \cap s \neq \emptyset$ and $B^c \cap s \neq \emptyset$ are not both true. Thus the four conditions for A and B to be qualitatively independent of each other in context s are satisfied. But now note that, by the definition of questions, and the fact that A and B are true throughout s , it follows that $A?_s = \{s, \emptyset\}$, and $B?_s = \{s, \emptyset\}$. By the definition of orthogonality, it follows that $A?_s$ is not orthogonal to $B?_s$, since there is an element of $A?_s$ (namely \emptyset) whose intersection with an element of $B?_s$ is \emptyset . And so, by the definition of question independence, we have that A and B are not question independent of each other in context s . Thus A and B are not question independent of each other in s , but are qualitatively independent of each other in s .

In his treatment of the proviso problem, van Rooij does not distinguish these two notions of independence, since he takes them to be equivalent. But since they are inequivalent, this raises an interpretive question: which notion of independence is the one van Rooij is arguing helps with the proviso problem? We think it is clear that it is qualitative independence, not question independence—and thus went this way in presenting van Rooij’s view in the main text. The reason for this is that if A and B are question independent in s , it follows that s does not entail any of A, B, A^c , or B^c . But then there is no way that the question independence of A and B in s can ever be part of an explanation of the fact that s entails B . And van Rooij’s use of independence is supposed to do just that: the independence of A and P in s , together with the assumption that A is compatible with s and that $A \supset P$ is entailed by s , is meant to show that s entails P . This will not follow if we interpret ‘independence’ as ‘question independence’, but it will follow if we interpret ‘independence’ as ‘qualitative independence’; and so we think the latter is the charitable interpretation of van Rooij’s main claims.

¹⁶‘Partition’ is sometimes defined in such a way that a partition can have no empty members; given this definition, van Rooij clearly has in mind a broader construal, on which partitions can have empty members.

B QUALITATIVE AND PROBABILISTIC INDEPENDENCE

Here we make clear the relation between qualitative and probabilistic independence. Let us restrict our attention to those models with finite outcome spaces, Ω , where the event space is simply the powerset of Ω , 2^Ω (call any such probability space a *finite probability space*). Such a model then can be described simply by an ordered pair $\langle W, p \rangle$ of a finite set W and a probability function p over the powerspace of W . Call the set of all such pairs P . Consider the natural mapping q from such pairs to sets of worlds: $q : P \rightarrow 2^W$, where $q(\langle W, p \rangle) = \{w \in W : p(\{w\}) > 0\}$.

Now we can state the relation between van Rooij's qualitative independence and probabilistic independence as follows:¹⁷

RELATION BETWEEN QUALITATIVE AND PROBABILISTIC INDEPENDENCE:
Given a context c and two propositions A and B , A is qualitatively independent of B with respect to c iff there is some finite probability space $\langle W, p \rangle$ such that (i) $q(\langle W, p \rangle) = c$ and (ii) A and B are probabilistically independent in $\langle W, p \rangle$.

Proof:

\Rightarrow Given qualitative independence for A and B relative to c , we show that there is a $\langle W, p \rangle$ which meets (i) and (ii). Note first that if c entails A or B or A^c or B^c , then $p(A)$ or $p(B)$ is 1 or 0, and so probabilistic independence is automatic. Now suppose otherwise; then we know that in c there are u $A \cap B$ worlds, x $A \cap B^c$ worlds, y $A^c \cap B$ worlds, and z $A^c \cap B^c$ worlds, for $u, x, y, z > 0$. Then suppose we have n probability mass. Let each world in the $A \cap B$ region receive $\frac{a}{u}$ probability mass; each world in the $A \cap B^c$ receive $\frac{b}{x}$ probability mass; each world in the $A^c \cap B$ region receive $\frac{c}{y}$ probability mass; and each world in the $A^c \cap B^c$ region receive $\frac{d}{z}$ probability mass, where a, b, c and d satisfy $a + b + c + d = n$ and $ad = bc$ (we could do this e.g. by setting all equal to $\frac{n}{4}$). Then we are guaranteed to have $P(A)P(B) = P(A \cap B)$.

\Leftarrow We show that for arbitrary probability space $\langle W, p \rangle$ such that A and B are probabilistically independent and $q(\langle W, p \rangle) = c$, qualitative independence holds for A and B with respect to c . Suppose that c includes A worlds and B worlds. Then $p(A)$ and $p(B)$ are greater than zero, but then $p(A \cap B) = p(A)p(B) > 0$, and so c includes $A \cap B$ worlds. Similar reasoning shows that, if c includes A and B^c worlds, it includes $A \cap B^c$ worlds, and so on for A^c and B , and for A^c and B^c ; crucial here is the fact that, if A and B are probabilistically independent, then so are A and B^c , A^c and B , and A^c and B^c .

¹⁷See Franke [2007, fn. 2] for a similar observation.

C UPDATING CONDITIONAL PROBABILITIES

Finally, we discuss two standard rules for updating conditional probabilities, and show that both systematically disrupt probabilistic independence. The update problem that Lassiter’s system gives us can be described as follows. Given a background probability function p_0 according to which A and B are independent, and non-maximal (i.e. not 0 or 1), and $p_0(B|A) < t$, how do we update p_0 to satisfy the condition that $p_0(B|A) \geq t$? Call the new function after the update p_1 . (For simplicity we will assume here that the update is minimal in the sense that we will have $p_1(B|A) = t$.)

The first method we consider, often called *infomin* (following van Fraassen 1981; see Kullback and Leibler 1951) is to minimize relative entropy, i.e. the Kullback-Leibler (KL) divergence. Relative entropy or KL divergence is a real number representing the ‘distance’ from one probability function to another.¹⁸ For two discrete probability functions p_0, p_1 over an outcome space A with finest partition E , we can define the KL divergence to p_1 from p_0 as follows:

$$\text{KL}(p_1, p_0) = \sum_{i \in E} p_1(i) \log \frac{p_1(i)}{p_0(i)}$$

Suppose there is some condition T on probability functions that p_0 does not meet. The infomin rule requires that the new probability function p_1 will be such that p_1 satisfies T , and for any other probability function p' that satisfies T , $\text{KL}(p', p_0) \geq \text{KL}(p_1, p_0)$. In other words, p_1 must be among the ‘closest’ probability functions that satisfy T , where closeness is measured by KL divergence.¹⁹

The second rule, from Douven and Romeijn [2011], minimizes not relative entropy but rather inverse relative entropy (IRE). For discrete probability functions p_0 and p_1 over outcome space A with finest partition E , IRE is defined as follows:

$$\text{IRE}(p_1, p_0) = \text{KL}(p_0, p_1) = \sum_{i \in E} p_0(i) \log \frac{p_0(i)}{p_1(i)}$$

Both of these rules have the following key property:

PRESERVATION OF CONDITIONAL PROBABILITIES: Updating a discrete probability function p to p' in order to change the conditional probability of B on A using either the rule *minimize KL* or the rule *minimize IRE* leaves unchanged the conditional probability of B on A^c , provided both p and p' assign non-maximal probability to A .

The proof is in a moment. First, note that this fact suffices to guarantee that, for any discrete probability function p , if p makes B and A probabilistically

¹⁸It is non-symmetric in that the distance from p_0 to p_1 is not always the same as the distance from p_1 to p_0 , hence it is not a measure.

¹⁹Our statement of the rule is non-deterministic in cases where there is no unique such function.

independent, then updating p by either rule to change the conditional probability of B on A will result in a probability distribution which does *not* make B and A probabilistically independent, provided both the original and updated functions assign A non-maximal probability. In other words, both rules under consideration are *guaranteed to disrupt* probabilistic independence in almost every case in which we use the rule to change a conditional probability of B on A . To see this, let $p(A) \neq 0$ and $p(A) \neq 1$ and let $p(B|A) = p(B) = t$. Suppose we update to a new probability distribution p' which meets the conditions that $p'(B|A) = k \neq t$ and $p'(A) \neq 0$ and $p'(A) \neq 1$ using one of these rules. Since $p(B|A) = p(B)$, it follows by the law of total probability that $p(B|A^c) = p(B)$; thus $p(B|A^c) = t$. Given PRESERVATION OF CONDITIONAL PROBABILITIES, we have $p'(B|A^c) = t$. By the law of total probability, we have $p'(B) = p'(B|A)p'(A) + p'(B|A^c)p'(A^c)$. Thus $p'(B) = kp'(A) + tp'(A^c)$. We will have $p'(B) = k$ iff $k = t$; but $k \neq t$; so $p'(B) \neq k$ so $p'(B) \neq p'(B|A)$: A and B are no longer probabilistically independent after our update.

This makes clear why, under either of these update rules, probabilistic independence of A and B will not be maintained when we update a probability distribution to increase the conditional probability of B on A ; it also makes clear how great of an assumption it would be to think that such updates generally do maintain probabilistic independences, and how much of a departure from standard ways of thinking about how such updates should go.

The proof of PRESERVATION OF CONDITIONAL PROBABILITIES is as follows.²⁰ Start with infomin. Consider any discrete probability functions p, p' over any event space with finest partition E . Let A be any union of members of E , i.e. any event, such that both p and p' assign non-maximal probability to A , and A^c its complement in the outcome space. First note:

$$\text{KL}(p', p) = \sum_{i \in E \& i \subset A} p'(i|A)p'(A) \log \frac{p'(i|A)p'(A)}{p(i|A)p(A)} + \sum_{i \in E \& i \subset A^c} p'(i|A^c)p'(A^c) \log \frac{p'(i|A^c)p'(A^c)}{p(i|A^c)p(A^c)}$$

Next note:

$$\begin{aligned} \sum_{i \in E \& i \subset A} p'(i|A)p'(A) \log \frac{p'(i|A)p'(A)}{p(i|A)p(A)} &= p'(A) \sum_{i \in E \& i \subset A} p'(i|A) \log \frac{p'(i|A)p'(A)}{p(i|A)p(A)} \\ &= p'(A) \sum_{i \in E \& i \subset A} p'(i|A) \left(\log \frac{p'(i|A)}{p(i|A)} + \log \frac{p'(A)}{p(A)} \right) \\ &= p'(A) \log \frac{p'(A)}{p(A)} + p'(A) \sum_{i \in E \& i \subset A} p'(i|A) \log \frac{p'(i|A)}{p(i|A)} \end{aligned}$$

²⁰Thanks to Gary Chamberlain for suggesting this proof method.

$$= p'(A) \log \frac{p'(A)}{p(A)} + p'(A) \text{KL}(p'_A, p_A)$$

where p_X is the probability function resulting from p conditioned on X . By like reasoning, we have

$$\sum_{i \in E \& i \subset A^c} p'(i|A^c) p'(A^c) \log \frac{p'(i|A^c) p'(A^c)}{p(i|A^c) p(A^c)} = p'(A^c) \log \frac{p'(A^c)}{p(A^c)} + p'(A^c) \text{KL}(p'_{A^c}, p_{A^c})$$

So we have:

$$\text{KL}(p', p) = p'(A) \log \frac{p'(A)}{p(A)} + p'(A) \text{KL}(p'_A, p_A) + p'(A^c) \log \frac{p'(A^c)}{p(A^c)} + p'(A^c) \text{KL}(p'_{A^c}, p_{A^c})$$

Obviously the last term is minimized by making $p'_{A^c} = p_{A^c}$, and thus by preserving all conditional probabilities on A^c . Since we can stipulate this without affecting the values of the other terms, we know that any function p' which minimizes KL from p will have this property. The proof for IRE is essentially identical; we show that:

$$\text{IRE}(p', p) = p(A) \log \frac{p(A)}{p'(A)} + p(A) \text{IRE}(p'_A, p_A) + p(A^c) \log \frac{p(A^c)}{p'(A^c)} + p(A^c) \text{IRE}(p'_{A^c}, p_{A^c})$$

One again, the last term, and the whole equation, is obviously minimized by making $p'_{A^c} = p_{A^c}$, and thus by preserving all conditional probabilities on A^c .²¹

REFERENCES

- David Beaver. *Presupposition and Assertion in Dynamic Semantics*. CSLI Publications: Stanford, CA, 2001.
- David I. Beaver. The kinematics of presupposition. In *ITLI Prepublication Series for Logic, Semantics and Philosophy of Language*. University of Amsterdam, 1992.
- Richard Bradley. Radical probabilism and Bayesian conditioning. *Philosophy of Science*, 72:342–64, 2005.
- C.E Alchourrón and P. Gärdenfors and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.
- Emmanuel Chemla. Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. Manuscript, Ecole Normale Supérieure, Paris and MIT, 2008.

²¹Douven and Romeijn [2011] also prove this fact about the minimize IRE rule using a different method.

- Igor Douven and Jan-Willem Romeijn. A new resolution of the Judy Benjamin problem. *Mind*, 120(479):637–670, 2011.
- Kai von Fintel. What is presupposition accommodation, again? *Philosophical Perspectives*, 22(1):137–170, 2008.
- Danny Fox. Presupposition projection from quantificational sentences: trivalence, local accommodation, and presupposition strengthening. Manuscript, 2012.
- Bas van Fraassen. Presuppositions, supervaluations and free logic. In Kirk Lambert, editor, *The Logical Way of Doing Things*, pages 67–92. Yale University Press, New Haven, 1969.
- Itamar Francez. Chimerical conditionals. *Semantics & Pragmatics*, 8(2):1–35, 2015.
- Michael Franke. The pragmatics of biscuit conditionals. In *Amsterdam Colloquium*, number 16, 2007.
- Gerald Gazdar. *Pragmatics: Implicature, Presupposition and Logical Form*. Academic Press, New York, 1979.
- Benjamin Ross George. Presupposition repairs: A static, trivalent approach to predicting projection. Master’s thesis, University of California, Los Angeles, 2008.
- Bart Geurts. Local satisfaction guaranteed: A presupposition theory and its problems. *Linguistics and Philosophy*, 19(3), 1996.
- Bart Geurts. *Presuppositions and Pronouns*. Elsevier, Amsterdam, 1999.
- Arno Goebel. Laws for biscuits: Law-like independence and biscuit conditionals. In *Semantics and Linguistic Theory (SALT)*, volume 27, 2017.
- Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst, 1982.
- Irene Heim. On the projection problem for presuppositions. In Michael Barlow, Daniel P. Flickinger, and Nancy Wiegand, editors, *The Second West Coast Conference on Formal Linguistics (WCCFL)*, volume 2, pages 114–125. Stanford, Stanford University Press, 1983.
- Irene Heim. Presupposition projection. In Rob van der Sandt, editor, *Presupposition, Lexical Meaning, and Discourse Processes: Workshop Reader*. University of Nijmegen, 1990.
- Irene Heim. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9(3):183–221, 1992.
- Irene Heim. On the proviso problem. Presentation to Milan Meeting, Gargnano, 2006.
- Hans Kamp. Presupposition computation and presupposition justification: One aspect of the interpretation of multi-sentence discourse. In Myriam Bras and Laure Vieu, editors, *Semantics and Pragmatic Issues in Discourse and Dialogue*, pages 57–84. Elsevier, Amsterdam, 2001.
- Lauri Karttunen. Presuppositions and linguistic context. *Theoretical Linguistics*, 1(1-3):181–194, 1974.

- Lauri Karttunen and Stanley Peters. Conventional implicatures in Montague grammar. In *Syntax and Semantics 11: Presupposition*. Academic Press, New York, 1979.
- Stephen Kleene. *Introduction to Metamathematics*. North-Holland, Amsterdam, 1952.
- S. Kullback and R.A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, 1951.
- Daniel Lassiter. Presuppositions, provisos, and probability. *Semantics and Pragmatics*, 5(2):1–37, 2012.
- Matthew Mandelkern. Dissatisfaction theory. In Mary Moroney, Carol-Rose Little, Jacob Collard, and Dan Burgdorf, editors, *Semantics and Linguistic Theory (SALT)*, volume 26, pages 391–416, 2016a.
- Matthew Mandelkern. A note on the architecture of presupposition. *Semantics and Pragmatics*, 9(13), 2016b.
- Stanley Peters. A truth-conditional formulation of Karttunen’s account of presupposition. In *Texas Linguistic Forum 6*. University of Texas at Austin, 1977.
- Robert van Rooij. Strengthening conditional presuppositions. *Journal of Semantics*, 24(3):289–304, 2007.
- Daniel Rothschild. Explaining presupposition projection with dynamic semantics. *Semantics and Pragmatics*, 4(3):1–43, 2011/2015.
- Rob van der Sandt. Presupposition and discourse structure. In Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, editors, *Semantics and Contextual Expression*, pages 267–294. Foris, Dordrecht, 1989.
- Rob van der Sandt. Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(3):333–377, 1992.
- Philippe Schlenker. Be articulate: a pragmatic theory of presupposition projection. *Theoretical Linguistics*, 34(3):157–212, 2008.
- Philippe Schlenker. Local contexts. *Semantics and Pragmatics*, 2(3):1–78, 2009.
- Philippe Schlenker. The proviso problem: a note. *Natural Language Semantics*, 19(4):395–422, 2011.
- Raj Singh. A solution to the proviso problem: Formally defined alternatives and relevance. Manuscript, Massachusetts Institute of Technology, 2006.
- Raj Singh. Formal alternatives as a solution to the proviso problem. In *Semantics and Linguistic Theory (SALT)*, volume 17, pages 264–281, 2007.
- Raj Singh. *Modularity and Locality in Interpretation*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 2008.
- Scott Soames. How presuppositions are inherited: a solution to the projection problem. *Linguistic Inquiry*, 13(3):483–545, 1982.
- Robert Stalnaker. Presuppositions. *Journal of Philosophical Logic*, 2:447–457, 1973.
- Robert Stalnaker. Pragmatic presuppositions. In Milton K. Munitz and Peter

Unger, editors, *Semantics and Philosophy*, pages 197–213. New York University Press, New York, 1974.

Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–86, 1975.

Peter Strawson. *Introduction to Logical Theory*. Methuen, London, 1952.

Bas van Fraassen. A problem for relative information minimizers in probability kinematics. *The British Journal for the Philosophy of Science*, 32(4):375–379, 1981.