



Robert van Rooij 

Patricia Mirabile 

Dynamic Probabilistic Entailment. Improving on Adams' Dynamic Entailment Relation

Abstract. The inferences of contraposition ($A \Rightarrow C \therefore \neg C \Rightarrow \neg A$), the hypothetical syllogism ($A \Rightarrow B, B \Rightarrow C \therefore A \Rightarrow C$), and others are widely seen as unacceptable for counterfactual conditionals. Adams convincingly argued, however, that these inferences are unacceptable for *indicative conditionals* as well. He argued that an indicative conditional of form $A \Rightarrow C$ has assertability conditions instead of truth conditions, and that their assertability 'goes with' the conditional probability $p(C|A)$. To account for inferences, Adams developed the notion of probabilistic entailment as an extension of classical entailment. This combined approach (correctly) predicts that contraposition and the hypothetical syllogism are invalid inferences. Perhaps less well-known, however, is that the approach also predicts that the unconditional counterparts of these inferences, e.g., modus tollens ($A \Rightarrow C, \neg C \therefore \neg A$), and iterated modus ponens ($A \Rightarrow B, B \Rightarrow C, A \therefore C$) are predicted to be valid. We will argue both by example and by calling to the results from a behavioral experiment ($N = 159$) that these latter predictions are incorrect *if the unconditional premises in these inferences are seen as new information*. Then we will discuss Adams' (1998) *dynamic probabilistic entailment relation*, and argue that it is problematic. Finally, it will be shown how his dynamic entailment relation can be improved such that the incongruence predicted by Adams' original system concerning conditionals and their unconditional counterparts are overcome. Finally, it will be argued that the idea behind this new notion of entailment is of more general relevance.

Keywords: probabilistic entailment; experimental philosophy

1. Introduction

There are four inference patterns that are valid for the material conditional but are crucially **invalid for natural language conditionals**:

1. **Contraposition:** $A \Rightarrow C \therefore \neg C \Rightarrow \neg A$.
2. **The hypothetical syllogism:** $A \Rightarrow B, B \Rightarrow C \therefore A \Rightarrow C$.
3. **Strengthening the antecedent:** $A \Rightarrow C \therefore (A \wedge B) \Rightarrow C$.
4. **Or-to-If:** $A \vee C \therefore \neg A \Rightarrow C$.

It is widely accepted that these inference patterns are, indeed, invalid for *counterfactual* conditionals (cf. Stalnaker, 1968; Lewis, 1973). However, Adams (1965, 1966, 1976), Cooper (1985) and others convincingly argued by examples that each of the above inference patterns should be invalid as well for indicative conditionals.¹ They argued on the basis of such examples that indicative conditionals should be analysed probabilistically: the assertability of $A \Rightarrow C$ is measured by the corresponding conditional probability of C given A , $p(C|A)$. Moreover, Adams (1975) showed that we can define a probabilistic entailment relation, \models^P , on the standard propositional language extended with the conditional connective ' \Rightarrow ' that allows for conditional sentences to have no truth conditions, which is a *conservative extension* of classical (propositional) logic. According to this new logic, none of the inference patterns listed above are valid, as Adams argued it should be. These same inference patterns also have non-conditional counterparts:

1. *Contraposition.* The non-conditional counterpart of contraposition is *modus tollens*: $A \Rightarrow C, \neg C \therefore \neg A$.
2. *The hypothetical syllogism.* The non-conditional counterpart is known as *iterated modus ponens*, and is given by $A \Rightarrow B, B \Rightarrow C, A \therefore C$.
3. *Strengthening the antecedent.* Its non-conditional counterpart, $A, B, A \Rightarrow C \therefore C$, has no standard name. It will be referred to here as *propositional premise monotonicity*.

¹ Stalnaker (1975) agrees, and argues that the fact that these inferences *seem reasonable* for indicative conditionals, or hold most of the time, can be accounted for by pragmatic means. Adams (1983) and Pearl (1990) proposed other pragmatic analyses to account for the same phenomenon. Notice that these pragmatic approaches all would strengthen probabilistic entailment, while we will argue that probabilistic entailment is already too strong. The issues are independent, however, and we agree that some such pragmatic approach should be pursued, but will ignore this issue in this paper.

4. *Or-to-If*. The non-conditional counterpart will be called *denying a disjunct*: $A \vee C, \neg A \therefore C$.

What is remarkable, however, is that whereas all of the first inference patterns are correctly predicted to be *invalid* by Adams' (1975) analysis, the non-conditional counterparts of these four patterns are all predicted to be *valid*. We discuss Adams' predictions in the next section before arguing, in Section 3, against the idea that a distinction should be made between the conditional inferences listed above and their non-conditional counterparts, *if the non-conditional premise is seen as new information*. We do this both with an argument by example and by reporting the results of an empirical study which indicates that the majority of reasoners reject the conclusions of those inference patterns, both in their conditional and non-conditional forms. In the fourth section we show that a slightly different entailment notion than \models^P , which we will call *dynamic probabilistic entailment* or \models^{dp} , allows the above incongruence to disappear: with \models^{dp} we predict that the conditional inferences and their non-conditional counterparts behave similarly in respect to their validity. In the concluding section we suggest that the idea behind the new dynamic entailment relation developed here is of more general relevance.

2. Adams' static notion of probabilistic entailment

For his definition of probabilistic entailment, Adams (1966, 1976) makes use of the notion of *uncertainty* given probability function p . The *uncertainty* of ϕ w.r.t. p , $U_p(\phi)$, is $1 - p(\phi)$. With this notion, Adams proves the following by now well-known facts for the language of propositional logic, \mathcal{L} (where \models^{cl} denotes the classical entailment relation):

FACT 1. *If $\Gamma \not\models^{cl} \phi$, then there is a p : $U_p(\phi) = 1$ and $\sum_{\gamma \in \Gamma} U_p(\gamma) = 0$.*

FACT 2. *$\Gamma \models^{cl} \phi$ iff for all p : $U_p(\phi) \leq \sum_{\gamma \in \Gamma} U_p(\gamma)$.*

Adams (1966, 1976) then defines a new notion of entailment, i.e. *probabilistic entailment*, \models^P , or *p-entailment*, as follows:

$$\Gamma \models^P \phi \quad \text{iff}_{df} \quad \text{for all } p: U_p(\phi) \leq \sum_{\gamma \in \Gamma} U_p(\gamma).$$

Given Fact 2, it follows immediately that for propositional language \mathcal{L} , \models^{cl} and \models^P coincide.

Adams defines the *conditional probability* of C given A , $p(C|A)$, as follows:

$$p(C|A) := \begin{cases} \frac{p(A \wedge C)}{p(A)} & \text{if } p(A) > 0, \\ 1 & \text{if } p(A) = 0. \end{cases}$$

Probabilistic entailment becomes interesting once we add the new conditional connective ' \Rightarrow ' to the language, \mathcal{L}^\Rightarrow , and assume that the assertability of indicative conditionals of the form $A \Rightarrow C$ should be measured by their corresponding conditional probabilities, $p(C|A)$. Language \mathcal{L}^\Rightarrow contains conditionals like $A \Rightarrow C$ only in case both A and C are formulae of the classical propositional language \mathcal{L} . Thus conditionals in this language do not 'embed'. Both for language \mathcal{L} and for language \mathcal{L}^\Rightarrow , the following fact holds:

FACT 3. $\Gamma \models^p \phi$ iff there is no probability function p in which all premises in Γ have probabilities arbitrary close to 1, while conclusion ϕ has a probability arbitrary close to 0.

Adams (1975) shows that, with this notion of entailment, many inferences involving indicative conditionals are valid that intuitively should be valid. Such inferences include, for instance, the following:

- **Identity:** $\models^p A \Rightarrow A$.
- **Modus Ponens:** $A, A \Rightarrow C \models^p C$.
- **Cautious Monotonicity:** $A \Rightarrow B, A \Rightarrow C \models^p (A \wedge B) \Rightarrow C$.

In contrast, and more to the point of this paper, the following inferences are (correctly) predicted to be **invalid** by Adams:

1. **Contraposition:** $A \Rightarrow C \not\models^p \neg C \Rightarrow \neg A$.
2. **The hypothetical syllogism:** $A \Rightarrow B, B \Rightarrow C \not\models^p A \Rightarrow C$.
3. **Strengthening the antecedent:** $A \Rightarrow C \not\models^p (A \wedge B) \Rightarrow C$.
4. **Or-to-If:** $A \vee C \not\models^p \neg A \Rightarrow C$.

In order to determine whether an inference is probabilistically valid, Adams uses the notion of *order of magnitude probability orderings* (abbreviated as *OMP-orderings*). OMP-orderings are linear orders between valuations, or worlds, such that $w_i \succ w_j$ means that the probability of valuation function, or world, w_i is an order of magnitude greater than the probability of valuation function, or world, w_j . A sentence then *holds* in such an ordering (where it is assumed that A and C are 'factual' sentences, i.e., members of \mathcal{L} , and therefore don't contain the connective ' \Rightarrow ') iff:

1. A holds in the ordering if A is true in the world that is highest in the ordering (denoted by w_1),
2. $A \Rightarrow C$ holds in the ordering if the highest world where $A \wedge C$ is true is higher than the highest world where $A \wedge \neg C$ is true (and, if $A \wedge \neg C$ is not true in any world then $A \Rightarrow C$ holds).

Once these definitions are established, the probabilistic validity of an inference can be determined using the following OMP-counterexample theorem, proven by Adams:

OMP-counterexample theorem. *If there is an OMP-ordering in which all of an inference's premises in Γ hold but its conclusion ϕ does not hold then its premises can have probabilities arbitrarily close to 1 while its conclusion has a probability arbitrarily close to 0, and thus $\Gamma \not\models^P \phi$. If there is no such ordering then it satisfies the generalized uncertainty sum condition and $\Gamma \models^P \phi$.*

The OMP-counterexample theorem also allows for an easy proof of a new theorem:

THEOREM 1 (Monotonicity and transitivity of probabilistic entailment). *For any $A_1, \dots, A_n, B_1, \dots, B_m$ and C in some conditionally extended language $\mathcal{L}^{\Rightarrow}$:*

1. *If $A_1, \dots, A_{n-1} \models^P C$, then $A_1, \dots, A_n \models^P C$;*
2. *If $A_1, \dots, A_n \models^P B_j$ for all $j \leq m$, and $B_1, \dots, B_m \models^P C$, then $A_1, \dots, A_n \models^P C$.*

PROOF. Let $\mathcal{L}^{\Rightarrow}$ be a conditionally extended language, and let $A_1, \dots, A_n, B_1, \dots, B_m$ and C be formulae in $\mathcal{L}^{\Rightarrow}$.

Ad 1. Suppose that A_1, \dots, A_{n-1} probabilistically entail C . Then any OMP-ordering in which A_1, \dots, A_{n-1} hold is one in which C also holds. Suppose that A_1, \dots, A_n hold in some OMP-ordering. Then C holds too, hence $A_1, \dots, A_n \models^P C$.

Ad 2. Suppose that $A_1, \dots, A_n \models^P B_j$ for all $j \leq m$, and $B_1, \dots, B_m \models^P C$. Now assume that A_1, \dots, A_n hold in some OMP-ordering \succ . Then by the OMP-counterexample theorem B_1, \dots, B_m hold in \succ as well, hence by the same theorem, C holds. So, again by the same theorem, $A_1, \dots, A_n \models^P C$. □

Note that in classical logic inferences can be valid in two ways, which are equivalent. Taking modus ponens as an example: modus ponens can be considered to be valid if for each set of premises Γ , if Γ entails $A \Rightarrow C$

and A , then it also entails C ; But modus can also considered to be valid if $A, A \Rightarrow C \therefore C$ holds. The following theorem shows that validity can also be interpreted in the same two equivalent ways in probabilistic logic:

THEOREM 2 (Valid entailment as indefeasibility). *Let A, B, C be formulae. Then: $\Gamma \models^p A$ and $\Gamma \models^p B$ imply $\Gamma \models^p C$ for all sets of premises Γ if and only if $A, B \models^p C$.*

PROOF. Suppose the former. Let $\Gamma = \{A, B\}$. Since $A, B \models^p A$ and $A, B \models^p B$ by monotonicity, $A, B \models^p C$. Hence, if $\Gamma \models^p A$ and $\Gamma \models^p B$ imply $\Gamma \models^p C$ for all sets of premises Γ , then $A, B \models^p C$.

Now suppose the latter. Let Γ be a set of premises such that $\Gamma \models^p A$ and $\Gamma \models^p B$, and let \succ be an OMP-ordering in which Γ hold. Then A and B hold in \succ , so by supposition C holds in \succ , hence $\Gamma \models^p C$. \dashv

These small theorems can now be used to prove that the following **non-conditional counterparts** of the above mentioned invalid inferences are predicted to be *p-valid* by Adams:

1. **Modus Tollens:** $A \Rightarrow C, \neg C \models^p \neg A$.
2. **Iterated modus ponens:** $A \Rightarrow B, B \Rightarrow C, A \models^p C$.
3. **Propositional premise monotonicity:** $A \Rightarrow C, A, B \models^p C$.
4. **Denying a disjunct:** $A \vee C, \neg A \models^p C$.

Starting with the proof of *p*-validity for *modus tollens*, let us assume that both premises hold in an ordering \succ . This means, (1), that $A \Rightarrow C$ holds, and thus that $A \wedge C \succ A \wedge \neg C$; and (2), that $\neg C$ also holds, i.e. that $\neg C$ is true in the highest world in the ordering. However, if A were to hold in this ordering, i.e., be true in the highest world in the ordering, then it would follow that $A \wedge C \not\succeq A \wedge \neg C$, which would contradict (1). Thus, we conclude that $\neg A$ is true instead in the highest world in the ordering, and thus that $A \Rightarrow C, \neg C \models^p \neg A$. The proof for *iterated modus ponens* follows immediately from theorem 1.2 together with the fact that modus ponens is \models^p -valid. Finally, the proof for *propositional premise monotonicity* follows immediately from theorem 1.1 and the proof for *denying a disjunct* follows from fact 2 in combination with Adams' definition of probabilistic entailment (given that ' \Rightarrow ' does not appear in the inference).

3. Inference patterns involving conditionals and their non-conditional counterparts

3.1. Inference patterns and static probabilistic entailment

In this section we argue that the predictions pointed out in the previous section are problematic **if the non-conditional premise(s)** of the non-conditional arguments *modus tollens*, *iterated modus ponens*, *propositional premise monotonicity* and *denying a disjunct* **are seen as new information**. In that case, we think that the inference patterns with conditionals and their non-conditional counterparts (those where the non-conditional premise is included as new information) should be analysed in the same way, and in particular that **these inferences should be determined to be invalid**.²

Contraposition and modus tollens. Consider the following example, which was also included in the experiment reported below: Lucy recently graduated from university, where she studied finance, and is now applying for jobs. She also currently lives at home with her parents in the suburbs. One of her parents, for example her mother, would in such a situation have good reason to take conditional (1a) to be plausible and acceptable. On the other hand, they would not necessarily believe that (1b) is true, because they can imagine other reasons why Lucy might still not be hired despite doing well in her interview (e.g., the interview could be very competitive). The existence of such counterexamples clearly points to the invalidity of the inference from (1a) to (1b) and suggests therefore that contraposition as an argument form should be invalid for indicative conditionals.

1. (a) If Lucy does well for her job interview at the bank, then she will be hired for the position.
- (b) If Lucy does not get hired for the job interview at the bank, then she did not do well in her job interview at the bank.

Intuitively, adapting this same example into the form of a modus tollens argument should not lead it to become acceptable, in particular

² The validity of *Propositional premise monotonicity* is especially problematic, because it means that one cannot even consistently express that $A, B, A \Rightarrow C, (A \wedge B) \Rightarrow \neg C$, although this seems to be possible. This is most easily illustrated by using generics, instead of conditionals: there is nothing wrong with saying that birds fly, that penguins (that are birds) exist, but that penguins do not fly.

because the same counterexamples will hold.³ Take, again, Lucy's situation, where Lucy's mother again believes (2a). Now suppose that Lucy's mother receives a new piece of information, namely that Lucy was told (by the interviewers at the bank, for instance) that she would not get the position. Lucy's mother will therefore come to believe that (2b), but that does not mean that she will now necessarily conclude that (2c), given that possible counterexamples readily come to mind.

2. (a) If Lucy does well for her job interview at the bank, then she will be hired for the position.
- (b) Lucy will not get hired for the position.
- (c) Lucy did not do well in her job interview at the bank.

The counterexample to contraposition can thus arguably also serve as a counterexample to modus tollens, if the non-conditional premise (2b) is seen as new information. To understand the importance of this condition, notice that for (2a)–(2c) to be an instance of invalid modus tollens, it is crucial that the corresponding conditional probability of (2a) be lower than 1, just as it should be lower than 1 for (1a) to let (1a)–(1b) be an invalid instance of contraposition.⁴

The hypothetical syllogism and iterated modus ponens. Consider once again the previous example, featuring Lucy and her job applications. Assuming the conditionals (3a) and (3b) appear to be acceptable (although again, and importantly, not certain) to Lucy's mother, she may

³ For further empirical evidence that modus tollens is seen as less natural than modus ponens, and is thus less likely to be considered 'valid' by reasoners (see Oaksford and Chater, 2008).

⁴ One reviewer argues that the idea that the inference from (2a) and (2b) to (2c) is not considered to be good does not show that the inference is invalid. The reviewer argues that the inference is bad, because it is *unsound*, rather than invalid; after (2b) is given, premise (2a) is no longer taken to be true. This might well be the case. At the same time, however, the reviewer takes the inference from (1a) to (1b) to be invalid, just like Adams did. But we wonder whether there is such a clear difference. Isn't it the case, for instance, that after the antecedent of (1b) is taken to be true (if only temporary, or for the sake of argumentation), premise 1a is not accepted anymore (if only temporary, or for the sake of argumentation). And why should it be the case that people reject argument (2a)–(2c) because it is (valid but) unsound? Why wouldn't we conclude from this, for instance, that the argument is valid, but that the conclusion would be inappropriate to assert? (because for an indicative conditional to assert appropriately, it has to be the case that the antecedent is a real-life possibility, with a proper non-zero probability)? We think that the most straightforward explanation why people reject an inference is that these people take the inference to be invalid.

still find the conclusion not (3c) acceptable. Here again, counterexamples to the argument are easy to come by (e.g., Lucy might prefer to remain at her parents' house even with her new position). This illustrates that there are cases where the hypothetical syllogism sometimes is *prima facie* fallacious.

3. (a) If Lucy does well in her job interview at the bank, then she will be hired for the position.
- (b) If Lucy is hired for the position, then she will move to live closer to the bank.
- (c) If Lucy does well in her job interview at the bank, then she will move to live closer to the bank.

Analogously to the case of contraposition and modus tollens, the hypothetical syllogism can be transformed into an iterated modus ponens with a non-conditional premise that is received as new information (i.e. Lucy telling her mother that she did well in her interview). Such an argument would be subject, again, to the same counterexamples as the hypothetical syllogism and would therefore strike one as being fallacious despite being valid according to static probabilistic entailment. And just as in the previous case, a necessary condition for these inferences to be invalid is that the conditional probability that corresponds with premise (3a) be smaller than 1, and that is indeed the case.

Strengthening the antecedent and propositional premise monotonicity.

The following two examples constitute cases where both strengthening the antecedent and propositional premise monotonicity can be shown to be logically fallacious:

4. (a) If Tweety is a bird, then he can fly. Therefore, if Tweety is a penguin (and thus a bird), he can fly.
- (b) If Tweety is a bird, then he can fly. Tweety is a penguin. Therefore, Tweety can fly.

The invalidity of these inference rests again on the fact that the conditional probability of the statement 'If Tweety is a bird, then he can fly' is high but still inferior to 1 because, of course, penguins are birds that cannot fly.

Or-to-If and denying a disjunct. Given that the climate in Amsterdam is very rainy and that snow in December is a regular occurrence, (5a) is acceptable. It does not, however, seem to justify inferring the conditional

(5b). Similarly, if one accepts proposition (6a) and then comes to accept that it won't rain in Amsterdam in December 2021, it still feels fallacious to infer (6c).

5. (a) Either it will rain or it will snow in Amsterdam in December 2021.
- (b) If it doesn't rain in Amsterdam in December 2021, it will snow.
6. (a) Either it will rain or it will snow in Amsterdam in December 2021.
- (b) It won't rain in Amsterdam in December 2021.
- (c) Therefore, it will snow in Amsterdam in December 2021.

Hence, denying a disjunct must be an invalid inference pattern, if read in a 'dynamic' way. It is crucial, again, for the probability of (5a) to be — though high — less than 1; the belief in the disjunction should be given up if it is assumed that the first disjunct is false. The same holds for the denying a disjunct inference.

3.2. Empirical support

In order to provide further support for the claim that the non-conditional inference patterns discussed above are invalid in the same way as the corresponding conditional inference patterns, we conducted a pre-registered behavioral experiment in which collected validity judgements by participants on inference patterns that were presented both in their conditional and non-conditional version.⁵ More specifically, the experiment focused on two of the inference patterns described above: contraposition and hypothetical syllogism. Following Adams, we predicted that participants would reject at a rate meaningfully above chance arguments presented in the contraposition and hypothetical syllogism forms. Contrary to Adams, however, we predicted that those same arguments, when presented instead in the form of their non-conditional counterparts, with the non-conditional premise being received as new information, would also be rejected at a rate meaningfully above chance by participants.

We collected data from 159 participants, who were assigned either to the contraposition or to the hypothetical syllogism condition. Participants were then shown three vignettes which briefly described a concrete, realistic situation (including the Lucy example discussed above) in which an agent holds a number of beliefs (i.e. the premises of the contraposition or hypothetical syllogism) and were asked to assess whether that same

⁵ Preregistration and supplementary online materials are available at <https://osf.io/nydw6/>.

agent could conclude to a new belief (i.e. the conclusion of the relevant argument). On the following page of the survey, participants were reminded of the vignette they had just read and were told that a new piece of information had been received by the agent (i.e. the non-conditional premise of the modus tollens or of iterated modus ponens argument, depending on the condition). In light of this new information, participants were asked again to assess whether the agent could now endorse a new belief (i.e., again, the conclusion of the relevant argument). Responses to the validity question were collected using a forced-choice paradigm with three options (and thus allowing for some grading): “Yes”, “No” and “I do not know”. For all experiments, whether in conditional or non-conditional form, the “I do not know”-answer was given in at most 10% of the cases. Because we were interested in whether participants would assess those arguments as *invalid* rather than valid, we chose to re-code validity answers, such that both “No” and “I do not know” answers would be counted as rejections (or at least non-adoption) of an argument.⁶

Our results reveal that between 70 and 87 percent of participants rejected the validity of the arguments they were shown. We further analysed these results using a Bayesian logistic regression model with weakly regularizing priors, which regressed the probability of an argument being rejected on the condition, the argument’s form and their interaction. Our model predicted all arguments as being likely to be rejected by reasoners at a rate reliably above chance. Figure 1 reports both the observed proportion of rejections and the estimated probability of rejection, with its 95 percent compatibility interval, for each type of argument. 95 percent compatibility intervals indicate the range of estimates that are, according to the statistical model, most compatible with the data that was used in the analysis.⁷

These findings support Adams’ predictions regarding the invalidity of two inference patterns that include indicative conditionals, i.e. contraposition and the hypothetical syllogism: not only are these arguments predicted as invalid by his notion of probabilistic entailment but they are also perceived as such by reasoners (and in particular non-logicians). However, these results also provide further support for our proposal that

⁶ But due to the small percentage of “I do not know”-answers, the results would not have changed enormously if we wouldn’t have done so.

⁷ For a more detailed discussion of the experimental procedure and of the statistical analyses, please refer to the Appendix.

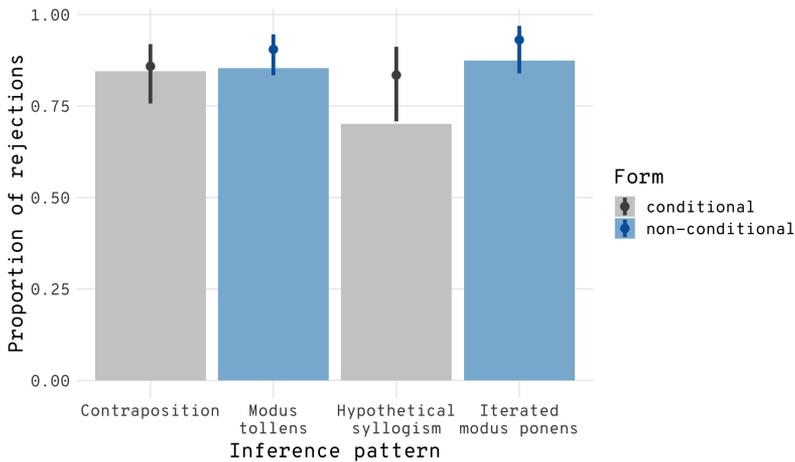


Figure 1. Proportion of arguments rejected as invalid by participants (bar plot) and mean estimated probability of rejection with 95 percent compatibility intervals.

the non-conditional inferences that were predicted by Adams to be valid should instead be analysed as invalid, and indeed, participants didn't think that the conclusion necessarily follows in a majority of cases.⁸

4. Dynamic probabilistic entailment

4.1. Adams' dynamic probabilistic entailment

Adams (1998) explicitly discusses examples in discourse of type (6a) and (6b) and notices that speakers generally do not accept both assertions. Instead, if (6a) was accepted first, the new information asserted by (6b)

⁸ Of course, it might well be that people reject an argument by contraposition because of invalidity and reject an argument by modus tollens because of unsoundness. As suggested by one reviewer, to separate these possibilities, we could ask participants in confidence judgements in both premises and conclusion after extra information is given (as in Over and Cruz, 2018). But as mentioned in an earlier footnote, we think that there are several other possibilities as well to 'explain' the data, and to separate them all we would have to do many experiments. That was not the point of this paper, and we prefer to interpret the data in what we take to be the most straightforward way: people reject both types of arguments because they take both arguments to be invalid.

leads reasoners to reconsider the belief in (6a).⁹ To account for the inferences that follow from (6a) and (6b), Adams suggests a new *dynamic* analysis of probabilistic entailment, which we might denote by \models^{da} , according to which premises and conclusion should be updated with the new information:

$$(7) \quad \Gamma, \phi_{\text{new}} \models^{\text{da}} \psi \quad \text{iff}_{\text{df}} \quad \text{for all } p: U_p(\psi|\phi) \leq \sum_{\gamma \in \Gamma} U_p(\gamma|\phi).$$

Unfortunately, Adams himself already remarks that this rule is problematic in several respects. First, his new entailment relation fails if one of the premises in Γ or the conclusion is a conditional of form $B \Rightarrow C$. Problems arise in this case first of all because Adams (1998) doesn't allow us to interpret $p(B \Rightarrow C|A)$. But even interpreting statements of that form as $p(C|A \wedge B)$, as could seem reasonable, gives rise to intuitively wrong predictions. In particular, Adams notices that it incorrectly predicts that $C, A_{\text{new}} \models \neg A \Rightarrow B$ becomes valid (because $p(B|A \wedge \neg A) = 1$), irrespective of the nature of C . A second problem noticed by Adams for his notion of dynamic entailment is that it doesn't make correct predictions when the new information is a conditional like $A \Rightarrow B$, which is due to the fact that Adams (1998) does not allow for conditional probabilities of the form $p(C|A \Rightarrow B)$ to be defined.¹⁰ Thus, Adams himself underlines that his proposal gives rise to various interpretability problems. For this reason, Adams' dynamic entailment relation does not make any predictions concerning modus tollens and iterated modus ponens.

Moreover, there is yet a further problem with Adams' dynamic entailment relation that is at least as serious: his analysis makes incorrect predictions even for those cases where it doesn't give rise to interpretability problems. In particular, it predicts that the denying a disjunct inference (6a)–(6c) is valid, although that should not be the case; or so we argued above.¹¹ As a consequence, Adams' dynamic notion of entailment contrasts with the natural idea that an agent accepts (or should accept)

⁹ Adams (1998) would say that the new factual premise is 'not in the scope of' at least one of the conditional premises. He also suggests the new premise in this case gives rise to *non-monotonic* behaviour of the inference-relation. He does not work out the latter suggestion, though.

¹⁰ There are ways to account for that, however, going back to ideas of de Finetti, as explained, for instance, by Milne (1997). We won't go into this here, as it is not our main point of criticism.

¹¹ To be sure, Adams (1998) is well aware that his notion of dynamic entailment makes inference (6a)–(6c) valid, and that this is counterintuitive. Instead of proposing an alternative notion of dynamic entailment to account for this intuition, his

an indicative conditional if and only if they accept the consequent after they learn the antecedent.

One of the most remarkable features of conditionals is that the strengthening of the antecedent inference is not valid because it shows that conditionals behave in a non-monotonic way. In the same way, the failure of propositional premise monotonicity is one of the defining features of non-monotonic logic: not only does the conditional behave non-monotonically but the inference relation does as well. In the next section we will define a dynamic inference relation that behaves non-monotonically exactly as conditionals do.

4.2. A new dynamic entailment relation

In this section, we propose an improvement on Adams' entailment relation \models^{da} that (i) is also dynamic, (ii) doesn't give rise to the problems that Adams' proposal does, and (iii) that predicts, in line with the results from section 3, a similar inferential behaviour for inference patterns with conditionals and inference patterns with their non-conditional counterparts. We denote this relation, \models^{dp} , as *dynamic probabilistic entailment*. Our new notion of entailment is strongly based on p -entailment but, to account for Adams' interpretability problems, we make a distinction between factual and conditional conclusions. To account for the similarity between the conditional and non-conditional inference patterns, we define an inference relation that is *non-monotonic*: $A_1, \dots, A_n \models^{\text{dp}} C$ does not imply $A_1, \dots, A_n, B \models^{\text{dp}} C$.

Our new entailment relation is a type of mixture of Adams' probabilistic entailment relation and the probabilistic analysis of *partial entailment* studied by Keynes (1921), Carnap (1950) and others. Partial entailment has been defined only for the language of classical logic, thus without sentences of the form ' $A \Rightarrow C$ ', such that ϕ follows to (at least) degree x ($0 \leq x \leq 1$) from a set of premises Γ , $\Gamma \models^x \phi$, iff the conditional probability of ϕ given the conjunction of the sentences in Γ is higher or equal than x , i.e., $p(\phi | \bigwedge_{\gamma \in \Gamma}) \geq x$. Students of partial entailment typically assumed that there is one 'objective' probability function, thus they did not have to quantify over all probability functions. But we might do

discussion focusses on the circumstances under which his dynamic entailment relation involving new information gives rise to counterintuitive predictions, i.e., when the new information is not 'in the scope of' the old premise(s). For a similar discussion from a more psychological point of view (see Oaksford and Chater, 2013).

so, as before. If so, it is not difficult to show that in case the premises in Γ are mutually consistent, $\Gamma \models^1 \phi$ iff $\Gamma \models^{\text{cl}} \phi$. Thus, partial entailment is a natural generalisation of classical entailment, but it is a very different generalisation of classical logic than Adams' generalisation. In particular, in contrast to Adams' static notion of entailment, partial entailment requires checking what happens if the (factual) premises are all assumed to be true (due to the use of conditionalization). That is, instead of considering the *prior probability* function, as Adams does, the *posterior probability* function needs to be considered instead. In this sense, partial entailment evaluates inferences from a *dynamic* point of view.

Our definition of a dynamic probabilistic entailment relation, \models^{dp} , handles 'old' premises as Adams does, but treats 'new' premises in the same way as is done in partial entailment, i.e. as certainties. To give a compact definition of this notion of *entailment*, let us first define another useful notion:

$$\Gamma \models_{\phi}^{\text{p}} \psi \text{ iff for all } p \text{ such that } p(\phi) = 1: U_p(\psi) \leq \sum_{\gamma \in \Gamma} U_p(\gamma).$$

Notice that if ϕ and ψ are factual, our $\Gamma \models_{\phi}^{\text{p}} \psi$ is very similar to Adams' $\Gamma, \phi_{\text{new}} \models_{\phi}^{\text{da}} \psi$. But Adams' entailment relation predicts wrongly for (6a)–(6c), or so we argued. To improve on \models^{da} , we define our analysis of *dynamic probabilistic entailment* as follows. Let $\Gamma \cup \{\phi, \psi\}$ be a (finite) set of formulae of some conditionally extended language $\mathcal{L}^{\Rightarrow}$. Then:

$$\Gamma, \phi_{\text{new}} \models^{\text{dp}} \psi \text{ iff}_{\text{df}} \begin{cases} \Gamma \models^{\text{p}} \phi \Rightarrow \psi & \text{if } \phi \text{ and } \psi \text{ are factual} \\ \Gamma \models_{\phi}^{\text{p}} \psi & \text{otherwise.} \end{cases}$$

We assume that if Δ is a set of factual sentences and ψ is factual, that $\Gamma, \Delta_{\text{new}} \models^{\text{dp}} \psi$ iff $\Gamma \models^{\text{p}} (\bigwedge \Delta) \Rightarrow \psi$, because Δ could equivalently have been given as one conjunction. Our new inference relation has some familiar and some unfamiliar properties. First, \models^{dp} is reflexive in that each premise A always entails itself: $\Gamma, A_{\text{new}} \models^{\text{dp}} A$. This is familiar. An unfamiliar property is that in contrast to \models^{p} , dynamic probabilistic entailment, \models^{dp} , is **non-monotonic**.

Example 1. Let A, C be factual formulae. Then $A \Rightarrow C, A_{\text{new}} \models^{\text{dp}} C$, since $A \Rightarrow C \models^{\text{p}} A \Rightarrow C$. Thus, if $\Gamma = \{A \Rightarrow C\}$, then $\Gamma, A_{\text{new}} \models^{\text{dp}} C$. Now take $\Gamma = \{A \Rightarrow C\}$ and $\Delta_{\text{new}} = \{A, \neg C\}$. Then $\Gamma, \Delta_{\text{new}} \models^{\text{dp}} A$, but $\Gamma, \Delta_{\text{new}} \not\models^{\text{dp}} C$, since $p(C|A \wedge \neg C) = 0$.

Entailment relation \models^{dp} is non-monotonic, in particular if the new information contradicts the earlier accepted premises. Example 1 more-

over shows not only that dynamic probabilistic entailment is defeasible but also that it makes even modus ponens defeasible. The example below shows this as well.

Example 2. Let Γ to be the following set of conditionals: $\Gamma = \{A \Rightarrow B, A \Rightarrow \neg C, B \Rightarrow C\}$, with A, B and C factual. Then it follows that $\Gamma, \{B\}_{\text{new}} \models^{\text{dp}} C$, but $\Gamma, \{A, B\}_{\text{new}} \models^{\text{dp}} \neg C$. (The latter follows by cautious monotonicity which is valid in \models^{p} and thus also in \models^{dp} .)

Although \models^{dp} is non-monotonic, its non-monotonic behavior is rather limited. In particular, in virtue of the monotonicity of \models^{p} , \models^{dp} is monotone if the conclusion is conditional:

THEOREM 3. *If $\Gamma \models^{\text{dp}} \phi$ and ϕ is conditional, then $\Gamma, \Delta_{\text{new}} \models^{\text{dp}} \phi$.*

5. Inference patterns in dynamic \models^{dp}

Because \models^{dp} is defined in terms of \models^{p} , and because we can use OMP-orderings to determine whether a p -entailment holds, it is relatively easy to determine \models^{dp} inference patterns as well if the new information and the conclusion are factual. Let us look at some examples.

One premise cases. Suppose $A \models^{\text{p}} C$, for factual A and C . Then also $(\top,) A_{\text{new}} \models^{\text{dp}} C$. This immediately follows from the fact that for factuais, $A \models^{\text{p}} C$ iff $A \models^{\text{cl}} C$, and that thus $p(C|A) = 1$ for each function p . Although $(A \Rightarrow C)_{\text{new}} \models^{\text{dp}} A \rightarrow C$ (with ‘ \rightarrow ’ material implication) just like for Adams’ static probabilistic entailment relation \models^{p} , it interestingly enough also follows that $(A \rightarrow C)_{\text{new}} \models^{\text{dp}} A \Rightarrow C$, which is in contrast to the behaviour w.r.t. \models^{p} . For the latter inference to be valid, it is crucial that we only look at probability functions that give $A \rightarrow C$ probability 1. Another interesting case is that although the paradox of material implication $C \not\models^{\text{p}} A \Rightarrow C$ is correctly predicted to be invalid probabilistically, it is predicted to be dynamically valid if C is seen as new information, $C_{\text{new}} \models^{\text{p}} A \Rightarrow C$, due to the fact that now C received probability 1.

From instant to conditional. First, note that $A, C \models^{\text{p}} A \Rightarrow C$. This is easily proven by the OMP-counterexample theorem. If \succ is some OMP-ordering that satisfies A and C , then A and C hold in the highest-ordered world w_1 . Then w_1 is also the highest-ordered world in which A holds, and as C holds in w_1 , this means that $A \Rightarrow C$ holds in \succ . From this

it follows immediately that $A, C_{\text{new}} \models^{\text{dp}} A \Rightarrow C$ and $(A \wedge C)_{\text{new}} \models^{\text{dp}} A \Rightarrow C$. Moreover, it holds that $A \wedge C, A_{\text{new}} \models^{\text{dp}} C$. More in general, it follows that in case χ follows classically from ϕ or ψ , and either $\phi \models^{\text{cl}} \psi$ or $\psi \models^{\text{cl}} \phi$, then $\phi, \psi_{\text{new}} \models^{\text{dp}} \chi$.

Basic modus ponens. Second, basic modus ponens is valid, i.e. $A, A \Rightarrow C \models^{\text{p}} C$. Again, by the OMP-counterexample theorem: if \succ is some OMP-ordering in which A holds, then $w_1 \models^{\text{cl}} A$. If $A \Rightarrow C$ also holds in \succ , then it must be that C is true in w_1 , hence C holds in \succ .

Moreover, $A \Rightarrow C, A \models^{\text{dp}} C$ if and only if $A \Rightarrow C \models^{\text{p}} A \Rightarrow C$, which is obviously valid. However, we have seen in the previous section that modus ponens is only a defeasible inference in dynamic probabilistic logic: it can be made defeated by adding a new premise.

The same argument can be given for the argument from $(A \wedge B) \Rightarrow C, B \models^{\text{p}} A \Rightarrow C$ and $(A \wedge B) \Rightarrow C, B_{\text{new}} \models^{\text{dp}} A \Rightarrow C$. Both are predicted to be valid.

Counterexample to modus ponens. (due to Cooper (1978)). Suppose that you believe $A =$ ‘These scales are accurate’ because somebody told that the statement was true. As new information, you see John, who looks like somebody of normal weight (B), get on the scale and you read the weight, which allows you to say $A \Rightarrow C =$ ‘If the scales are accurate, then John weighs four hundred points’. In such a situation, you would hardly conclude that $C =$ ‘John weighs four hundred pounds’. The reason is that $p(C|B)$ is taken to be very low. Indeed, $A, \{B, A \Rightarrow C\}_{\text{new}} \not\models^{\text{dp}} C$. In our case, $\{B, A \Rightarrow C\}_{\text{new}}$ gives us reason to give up on A .

From counterexample to no conditional. A and $\neg C$ together contradict $A \Rightarrow C$: $A, \neg C, A \Rightarrow C \models^{\text{p}} C$, by the monotonicity of entailment, while also $A, \neg C, A \Rightarrow C \models^{\text{p}} \neg C$ by monotonicity. Hence $A, \neg C, A \Rightarrow C \models^{\text{p}} \perp$.

However, one cannot validly reason ad absurdum from an indicative conditional and its counterexample in dynamic probabilistic logic: $A \Rightarrow C, \{A, \neg C\}_{\text{new}} \models^{\text{dp}} \perp$ if and only if $A \Rightarrow C \models^{\text{p}} (A \wedge \neg C) \Rightarrow \perp$. Again, any choice of A and C where C is not classically entailed by A suffices as a counterexample.

We can also consider inference patterns that are classically valid but probabilistically invalid, and where this invalidity is justified by counterexamples. Notable among them, as discussed in earlier sections, are **contraposition**, **the hypothetical syllogism**, and **strengthening of the antecedent**. Notice that for none of those inferences are factual sentences

involved, which means that these patterns hold in dynamic probabilistic logic if and only if they hold in (static) probabilistic logic. Thus **none of these inferences are valid** according to \models^{dp} .

Earlier we noted that the non-conditional variants of those inferences were still valid with respect to \models^{p} , and we argued that this was problematic. Fortunately, our new dynamic inference relation \models^{dp} does much better, and predicts all three inferences to be **invalid**:

Modus tollens. We showed above that $A \Rightarrow C, \neg C \models^{\text{p}} \neg A$. However, modus tollens is invalid in dynamic probabilistic logic. $A \Rightarrow C, \neg C_{\text{new}} \models^{\text{dp}} \neg A$ if and only if $A \Rightarrow C \models^{\text{p}} \neg C \Rightarrow \neg A$, i.e. if and only if contraposition is probabilistically valid for A and C . But the latter is not the case. Thus, $A \Rightarrow C, \neg C_{\text{new}} \not\models^{\text{dp}} \neg A$.¹²

Iterated modus ponens. In general, the hypothetical syllogism is invalid in probabilistic logic. That is: $A \Rightarrow B, B \Rightarrow C \not\models^{\text{p}} A \Rightarrow C$. This is illustrated by an OMP-counterexample in which $w_1 \models \neg A, B, C$ and $w_2 \models A, B, \neg C$. However, the hypothetical syllogism pattern *can* be used to infer C from A , since $A, A \Rightarrow B, B \Rightarrow C \models^{\text{p}} C$. Let \succ be an OMP-ordering in which the premises $A, A \Rightarrow B$ and $B \Rightarrow C$ all hold. Then $w_1 \models A$, so $w_1 \models B$. Hence $w_1 \models C$, so C holds in \succ .

Because, $A \Rightarrow B$ and $B \Rightarrow C$ entail $A \Rightarrow C$ in dynamic probabilistic logic only if they do so in probabilistic logic, it follows, however, that $A \Rightarrow B, B \Rightarrow C, A_{\text{new}} \not\models^{\text{dp}} C$.

Propositional premise monotonicity. Consider an OMP-ordering \succ , now such that $w_1 \models A, \neg B, C$, and $w_2 \models A, B, \neg C$. Then \succ is an OMP-counterexample to the pattern of *strengthening the antecedent*, or in other words, it shows that $A \Rightarrow C \not\models^{\text{p}} (A \wedge B) \Rightarrow C$. However, $A \wedge B, A \Rightarrow C \models^{\text{p}} A$, so that by transitivity, $A \wedge B, A \Rightarrow C \models^{\text{p}} C$.

Propositional premise monotonicity holds in dynamic probabilistic logic if and only if strengthening of the antecedents holds for its conditional counterpart. Thus, $A \Rightarrow C, (A \wedge B)_{\text{new}} \not\models^{\text{dp}} C$.

Finally, let us consider our last example where Adams' static probabilistic and dynamic entailment relations seem to seem to make incorrect predictions, i.e., the non-conditional variant of **from Or-to-If**.

Denying a disjunct. The Or-to-If pattern consists in inferring $A \Rightarrow C$ from $\neg A \vee C$, and is probabilistically invalid, as is shown by the

¹² Note that although $A \Rightarrow C, \neg C_{\text{new}} \not\models^{\text{dp}} \neg A$, it is the case that $\{A \Rightarrow C, \neg C\}_{\text{new}} \models^{\text{dp}} \neg A$.

OMP-counterexample \succ in which $w_1 \models \neg A, C$ and $w_2 \models A, \neg C$. Thus $\neg A \vee C \not\models^P A \Rightarrow C$.

But for exactly the same reason $\neg A \vee C, A_{\text{new}} \not\models^{\text{dp}} C$. Thus, denying a disjunct is **invalid** in dynamic probability logic.

6. Conclusion

In this paper, we have proposed that Adams' (1975) observation that the inference patterns *contraposition*, *the hypothetical syllogism*, *strengthening the antecedent* and *from or to if* are invalid for indicative conditionals (if analysed as conditional probabilities) needs to be extended to the case of the non-conditional counterparts of the inference patterns. Adams' probabilistic entailment correctly accounts for the intuitive invalidity of the first four inferences but it also predicts that the latter four inferences are valid. We argued that Adams' predictions regarding the non-conditional inferences listed above are incorrect on the basis of an argument by cases and we also report empirical results that indicate that these predictions contradict the intuitions of laypeople. Finally, we have shown that this false prediction can be avoided by using a *dynamic probabilistic entailment* relation which treats factual premises differently from conditional premises.

The idea behind our dynamic probabilistic entailment relation has a more general relevance. In section 3, we reminded the reader that Adams' notion of probabilistic entailment can also be captured in terms of order of magnitude probability orderings. It is well-known that such orderings closely correspond with Lewis's (1973) (and Stalnaker's (1968)) *qualitative* possibility orderings used for the analysis of counterfactuals, and more directly with the orderings that characterise valid inferences in the minimal non-monotonic logic: system **P** (see, e.g., Burgess, 1981).

To see this, we start with a preorder \succeq on possible worlds (a binary relation that is reflexive and transitive).¹³ As above, we can define in terms of this (plausibility) ordering an ordering between *sets* of possible worlds. If X and Y are sets of possible worlds, we say that $X \succ Y$ if and only if there is an $x \in X$ such that for any $y \in Y$ we have $x \succ y$, where $x \succ y$ iff_{df} $x \succeq y$ and $y \not\succeq x$ (note that \succ is thus a strict partial order, an order where \succ is irreflexive and transitive). In

¹³ The order \succeq on X is reflexive iff for any $x \in X$ we have $x \succeq x$. The order is transitive iff for all $x, y, z \in X$ if $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

terms of these orderings, we can say that $X \Rightarrow Y$ *holds* in \succeq , or is *accepted* in \succeq , iff $(X \wedge Y) \succ (X \wedge \neg Y)$. Similarly, X is accepted in \succ iff $(\top \Rightarrow X) \succ (\top \Rightarrow \neg X)$. Logical consequence can now be defined in an almost standard way, but now as *maintenance of acceptance* (Burgess, 1981). Just like the probabilistic entailment relation \models^P , this system invalidates the inferences *contraposition*, *hypothetical syllogism*, *strengthening the antecedent* and *from or to if*, but it validates the non-conditional counterparts of these inferences. Thus, the problem that we identified for the probabilistic analysis in section 2 of this paper is a more general problem, and holds for qualitative variants of Adams' analysis of conditional reasoning as well.

Fortunately, the solution we proposed to solve the problem does not make essential use of probabilities either. In fact, our definition of entailment can be used for qualitative versions of conditional reasoning as well, and with the same desirable effects: it would invalidate the non-conditional versions of the inferences *contraposition*, *hypothetical syllogism*, *strengthening the antecedent* and *from or to if*, just as it should. More generally, our approach is in line with the view that an agent accepts (or should accept) an indicative conditional if and only if they accept the consequent after they learn the antecedent.

A. Appendix: Additional experimental methods and results

A.1. Participants

We sought to detect a probability of argument rejection in any condition that would be above 65 percent, i.e. that would be meaningfully above chance. A simulation-based power analysis indicated that with a sample of 80 participants per condition, effects equal or larger than 65 percent would be successfully detected in 96 percent of cases. After accounting for an expected drop-out rate of approximately 15 percent, we recruited 182 participants, of which 180 gave complete response sets, and assigned them randomly to the contraposition (CP) or hypothetical syllogism (HS) conditions. The experiment was distributed via the Prolific.co platform. Participants were required to have English as a first language and to not have received a dyslexia diagnosis. After excluding, as was planned in our preregistration, participants who reported having received advanced training in logic, as well as participants who failed an

attention check or any of the two manipulation checks, we were left with data from 159 participants (136 women, 23 men; $M_{\text{age}} = 37.7$, $SD_{\text{age}} = 12.6$), with 82 participants in the contraposition condition and 77 participants in the hypothetical syllogism condition. Participants received should be 1.04 £ as compensation.

A.2. Materials and procedure

In addition to the validity task described in the main body of the paper, participants also completed a counterexample generation task immediately after the first validity task for each vignette. In the counterexample generation task, participants were prompted to attempt to generate a counterexample for the argument they had just considered. More specifically, they were asked whether they could “think of a reason why [the agent] could still NOT want to conclude” to the conclusion of the inference pattern. Participants were instructed to only generate one counterexample and were asked to indicate ‘No reason’ if they could not find a counterexample. The experiment also included, at fixed intervals, two comprehension checks in which participants were asked to perform the validity and the counterexample generation tasks for two modus ponens arguments. For these two modus ponens arguments, we expected participants to respond ‘Yes’ to the validity task. Participants who responded otherwise were scored as having failed the comprehension checks.

A.3. Analytic approach and results

To examine our predictions regarding the probability of the different argument forms being rejected, we fit a Bayesian regression model with the R package `brms` (Bürkner, 2018) and the probabilistic language Stan (Carpenter et al., 2017), which uses Markov Chain Monte Carlo algorithms. A Bayesian analysis estimates model parameters (in this case, the probability of rejection for the different arguments) as probability distributions, with the joint probability distribution of the data, y , and a given parameter, θ , being computed via the prior probability of θ and the probability $p(y|\theta)$:

$$p(y, \theta) = p(y|\theta) \times p(\theta).$$

This result is derived from Bayes’ Rule, which serves to calculate the posterior probability, $p(\theta|y)$:

$$p(\theta|y) \propto p(y|\theta) \times p(\theta) = p(y, \theta).$$

This posterior probability distribution can be interpreted as indicating the relative plausibility of possible values of the parameter θ , conditional on the prior probability of that parameter, the probability distribution of the responses (or likelihood function), and the data itself.

Because the response variable we sought to model was binary, we chose to model it as arising from a Bernoulli distribution. This model estimated the logit-transformed probability of an argument being rejected given the condition and form it was shown in. The logit-transformation converts a probability p (which is, by definition, restricted to the 0 to 1 range) into a log odds ratio by taking the logarithm of the ratio between p and $1-p$. A log odds ratio of 0 means that p and $1-p$ are independent, a positive log odds ratio means that p is higher than $1-p$, and a negative log odds ratio means that p is lower than $1-p$.

We also specified prior distributions over the possible probability estimates for each condition.¹⁴ Specifying these priors is recommended because it allows regularization of parameter estimates (see, for instance, Bürkner, 2018; McElreath, 2020). The priors we specified indicated that the probability that arguments shown in their non-conditional form would be rejected was likely to range between 11 and 88 percent with a mean of 50 percent, which corresponds to a wide prior that indicates extreme probabilities of rejection as unlikely while remaining agnostic to the direction of the effect. However, because of the already existing theoretical and empirical consensus over the invalidity of these arguments in their conditional form, we indicated that the probability that arguments shown in their conditional form would be rejected ranged between 37 and 92 percent with a mean of 73 percent, which corresponds to a narrower prior compatible with higher probability estimates. Finally, because we used a repeated measures design, where participants evaluated multiple vignettes and where vignettes were rated by multiple participants, we also included a (hierarchical) mixed-effects structure in our model, which estimates how group-level (or random) effects deviate from population-level (or main) effects and accounts for possible correlations in responses provided by the same participant or to the same vignette. Figure 2 represents the prior distributions along with the posterior distributions of estimated probabilities for Model 1. Our model was defined as follows:

¹⁴ Here, we deviated from our pre-registered priors in order to correct a misunderstanding about the structure of our model. These updated priors did not qualitatively modify the results of this analysis.

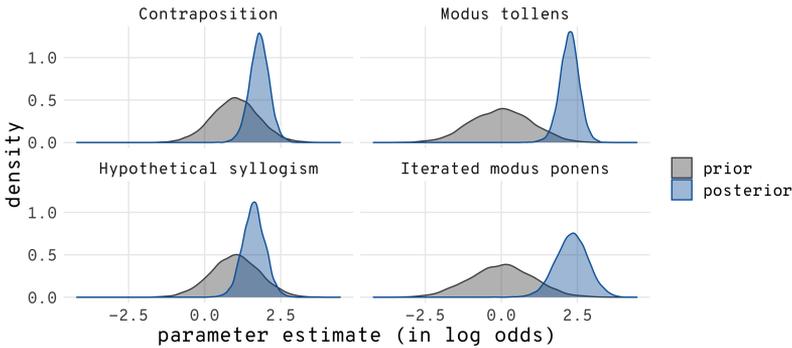


Figure 2. Prior distribution of model parameters (black dashed line) and corresponding posterior distributions.

Model 1: $\text{Rejection} \sim 0 + \text{Form} * \text{Condition} +$
 $(0 + \text{Form} + \text{Condition} \mid \text{Participant}) +$
 $(0 + \text{Form} + \text{Condition} \mid \text{Vignette})$ ¹⁵

We fitted the model with 4 chains and 4000 samples (2000 post warm-up) per chain. MCMC diagnostics indicated sufficient mixing of the chains, sufficiently high bulk and tail effective sample size values and an \hat{R} convergence diagnostic of 1.00 for all parameters, which is within the recommended value range (Vehtari, Gelman, Simpson, Carpenter, & Bürkner, 2020). This model estimated a positive log odds (i.e., a probability above chance) for all argument types: $b = 1.80$, [1.13 : 2.43] 95% CI for the contraposition, $b = 2.25$, [1.61 : 2.86] 95% CI for the modus tollens, $b = 1.61$, [0.88 : 2.34] 95% CI for the hypothetical syllogism and $b = 2.31$, [1.24 : 3.34] 95% CI for the iterated modus ponens.

Finally, we cross-validated our model using the `loo` package (Vehtari, Gabry et al., 2020) which performs leave-one-out cross-validation (Vehtari, Gelman & Gabry, 2017). This method provides estimates of the point-wise out-of-sample prediction accuracy (or ELPD) of a model, as well as approximations of the standard error for the estimated prediction error of a model, thereby enabling comparisons in predictive accuracy between models. We cross-validated Model 1 by comparing it with three restricted models: a model that only included Condition as a predictor (Model C), a model that only included Form as a predictor (Model F)

¹⁵ The 0 + syntax indicates that no separate intercept was estimated for this model.

Model	Δ_{ELPD}	(SE)
0	0.0	0.0
1	-0.6	7.2
F	-8.0	6.6
C	-118.5	14.2

Note: For each cross-validation comparison, the Δ_{ELPD} is the difference in ELPD between each model and the model with the largest ELPD (indicated in bold face). The Δ_{ELPD} for the best model is always 0 given that it is the difference between that model and itself. SE corresponds to the standard error of the difference in ELPD between two models and indicates the amount of uncertainty present in the comparison. The Δ_{ELPD} must be several times larger than the SE for the comparison to indicate meaningful differences in predictive performance between two models.

Table 1. Cross-validation comparisons of the models.

and a null-model (Model 0). The difference in ELPD between Model C and Model 0 (the best model) is multiple times larger (namely 8 times larger) than the standard error of that difference, which indicates that a null model has better predictive performance than a model that only considers the condition an argument was in. However, the difference in ELPD between Model 0 and Models 1 and F was either smaller or less than twice larger than the standard error of that difference (see Table 1 for results). This indicates that including Form, either by itself or with Condition and their interaction, as a predictor did not meaningfully improve the accuracy of predictions. These results can be interpreted as meaning that the accuracy in predicting the probability of an argument being rejected did not improve when the type of argument was taken into account. This suggests that participants in this experiment rejected the validity of the arguments they were shown in a way that was not meaningfully sensitive to the type of argument they were shown.

A.4. Counterexample generation task

The results from the counterexample generation task revealed that only in 6 percent of cases where participants indicated an argument as not valid did they also indicate being unable to generate a counterexample. This further suggests that counterexamples were readily available to participants who did not believe that the conclusion should follow from the available premises.

Acknowledgments. We would like to thank the reviewers of this paper for their comments. The work for this paper was supported by NWO grant “From Learning to Meaning: A new approach to Generic Sentences and Implicit Biases” (dossiernummer 406.18.TW.007).

References

- Adams, E., 1965, “The logic of conditionals”, *Inquiry* 8: 166–197. DOI: [10.1080/00201746508601430](https://doi.org/10.1080/00201746508601430)
- Adams, E., 1966, “Probability and the logic of conditionals”, pages 265–316 in J. Hintikka and P. Suppes (eds.), *Aspects of Inductive Logic*, North Holland Publishing Company, Amsterdam. DOI: [10.1016/S0049-237X\(08\)71673-2](https://doi.org/10.1016/S0049-237X(08)71673-2)
- Adams, E., 1975, *The Logic of Conditionals. An Application of Probability to Deductive Logic*, Reidel, Dordrecht.
- Adams, E., 1983, “Probabilistic enthymemes”, *Journal of Pragmatics* 7: 283–295.
- Adams, E., 1998, *A Primer of Probability Logic*, CSLI Publications, Stanford.
- Burgess, J. P., 1981, “Quick completeness proofs for some logics of conditionals”, *Notre Dame Journal of Formal Logic* 22: 76–84. DOI: [10.1305/nd-jfl/1093883341](https://doi.org/10.1305/nd-jfl/1093883341)
- Bürkner, P.-C., 2018, “Advanced Bayesian multilevel modeling with the R package brms”, *The R. Journal* 10 (1): 395–411. DOI: [10.32614/RJ-2018-01](https://doi.org/10.32614/RJ-2018-01)
- Carnap, R., 1950, *Logical Foundations of Probability*, Routledge and Kegan Paul, London. DOI: [10.2307/2021419](https://doi.org/10.2307/2021419)
- Carpenter, B., A. Gelman, M. Hoffman, D. Lee, B. Goodrich, B. Betancourt et al., 2017, “Stan: A probabilistic programming language”, *Journal of Statistical Software, Articles* 76 (1): 1–32. DOI: [10.18637/jss.v076.i01](https://doi.org/10.18637/jss.v076.i01)
- Cooper, W. S., 1978, *Foundations of Logico-Linguistics. A Unified Theory of Information, Language and Logic*, Reidel, Dordrecht.
- Keynes, J. M., 1921, *A Treatise on Probability*, Macmillan, London.
- Lewis, D., 1973, *Counterfactuals*, Basic Blackwell, Oxford. DOI: [10.2307/2273738](https://doi.org/10.2307/2273738)
- McElreath, R., 2020, *Statistical Rethinking. A Bayesian Course with Examples in R and Stan*, CRC Press.

- Milne, P., 1997, “Bruno de Finetti and the logic of conditional events”, *British Journal of Philosophy of Science*, 48: 195–232. DOI: [10.1093/bjps/48.2.195](https://doi.org/10.1093/bjps/48.2.195)
- Oaksford, M., and N. Chater, 2008, “Probability logic and the modus ponens: Modus tollens asymmetry in conditional inference”, pages 97–120 in M. Oaksford and N. Chater (eds.), *The Probabilistic Mind. Prospects for Bayesian Cognitive Science*, Oxford: Oxford University Press. DOI: [10.1093/acprof:oso/9780199216093.003.0005](https://doi.org/10.1093/acprof:oso/9780199216093.003.0005)
- Oaksford, M., and N. Chater, 2013, “Dynamic inference and everyday conditional reasoning in the new paradigm”, *Thinking & Reasoning* 19: 346–379. DOI: [10.1080/13546783.2013.808163](https://doi.org/10.1080/13546783.2013.808163)
- Over, D.E., and N. Cruz, 2018, “Probabilistic accounts of conditional reasoning”, pages 434–450 in L. Ball and V. Thompson (eds.), *International Handbook of Thinking and Reasoning*, Hove: Psychology Press.
- Pearl, J., 1990, “System Z: A natural ordering of defaults with tractable applications to nonmonotonic reasoning”, pages 121–135 in *TARK’90: Proceedings of the 3rd conference on Theoretical aspects of reasoning about knowledge*.
- Stalnaker, R., 1968, “A theory of conditionals”, in N. Rescher (ed.), *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, No. 2, Basil Blackwell, Oxford.
- Stalnaker, R., 1975, “Indicative conditionals”, *Philosophia* 5: 269–286. DOI: [10.1007/BF02379021](https://doi.org/10.1007/BF02379021)
- Vehtari, A., J. Gabry, M. Magnusson, Y. Yao, P. Bürkner, T. Paananen, and A. Gelman, 2020, *loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models*. DOI: [10.1007/s11222-016-9696-4](https://doi.org/10.1007/s11222-016-9696-4)
- Vehtari, A., A. Gelman, and J. Gabry, 2017, “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC”, *Statistics and Computing* 27 (5): 1413–1432. DOI: [10.1007/s11222-016-9696-4](https://doi.org/10.1007/s11222-016-9696-4)
- Vehtari, A., A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner, 2020, “Rank-normalization, folding, and localization: An improved \hat{R} for assessing convergence of MCMC”, *Bayesian Analysis*. DOI: [10.1214/20-BA1221](https://doi.org/10.1214/20-BA1221)

ROBERT VAN ROOIJ and PATRICIA MIRABILE
 Institute for Logic, Language and Computation
 University of Amsterdam, the Netherlands
 R.A.M.vanRooij@uva.nl
 p.l.mirabile@uva.nl