‘Might’ as a generator of alternatives: the view from reasoning *

Salvador Mascarenhas
Ecole Normale Supérieure
Department of Cognitive Studies
Institut Jean-Nicod
ENS, EHESS, PSL University
Paris France

Léo Picat
Ecole Normale Supérieure
Department of Cognitive Studies
Institut Jean-Nicod
ENS, EHESS, PSL University
Paris France

Abstract We argue that the epistemic modal ‘might’ is a generator of alternatives in the sense of Hamblin semantics (Kratzer & Shimoyama 2002) or inquisitive semantics (Ciardelli, Groenendijk & Roelofsen 2009). Building on methodologies from the psychology of reasoning, we show that ‘might’ patterns with disjunctions and with indefinites in giving rise to a particular kind of illusory inference. The best extant accounts of these illusory inferences crucially involve alternatives, paired with matching strategies (Walsh & Johnson-Laird 2004) or with question-answer dynamics (Koralus & Mascarenhas 2013). Our results constitute further steps toward convergence between theories and methodologies in natural language semantics and the psychology of reasoning.

Keywords: weak epistemic modals, illusory inferences, reasoning, mental models

1 Introduction

The semantics of epistemic modals has puzzled linguists and philosophers for decades. Here we address a debate that hasn’t been under the spotlight in recent years: the role of ‘might’ as a means of directing hearer attention by generating a single alternative in the sense of Hamblin semantics or inquisitive semantics.

Disjunction and indefinites are generators of alternatives par excellence. They induce a particular kind of illusory inference whose extant accounts all agree must be due to the presence of alternatives. We show experimentally that ‘might’ induces the same kind of illusory inference. Our results provide arguments from the psychology of reasoning in favor of particular theories of ‘might’.

* We thank Benjamin Spector and the audience of WIP-LANG at Ecole Normale Supérieure for extremely useful comments. The work presented here was funded by Agence Nationale de la Recherche grants ANR-17-EURE-0017 (Frontiers of Cognition) and ANR-18-CE28-0008 (Between Language and Reasoning).

©2019 Mascarenhas and Picat
In the remainder of this introduction, we present the minimal required background knowledge from the semantics literature on alternatives, the relevant theories from the psychology of reasoning, and the central empirical discoveries about the kinds of illusory inferences of interest. Section 2 presents our study and results, and section 3 examines the consequences of our results for a range of theories of the semantics of ‘might’. Section 4 concludes with remarks on ongoing work.

1.1 Illusory inferences from disjunction

Recent efforts to seek convergence between natural language semantics and the psychology of reasoning have led to articulated theories of interpretive processes and general-purpose reasoning. In particular, the erotetic theory of reasoning of Mascarenhas and Picat (2013) incorporates linguistic insights into a variant of mental models theory of reasoning (Johnson-Laird 1983) to account for a class of attractive fallacies known as illusory inferences from disjunction. Consider the example in (1), heavily simplified from the original examples discovered by Walsh & Johnson-Laird (2004).

(1) John speaks English and Mary speaks French, or else Bill speaks German.
John speaks English.
**Fallacious conclusion**: Mary speaks French.

The problem in (1), as well as multiple structurally identical problems, has an acceptance rate between 80% and 85% (Walsh & Johnson-Laird 2004; Mascarenhas & Koralus 2017; Koralus & Mascarenhas 2018). Yet, the inference in (1) is a fallacy. Suppose Bill speaks German (modeling premise 1) and John speaks English (premise 2), but Mary does not speak French. This is a model of the premises but not the conclusion, and the inference in (1) is thus invalid.

1.2 Mental models and illusory inferences

Building on the mental-models explanation of these fallacies by Walsh & Johnson-Laird (2004), the erotetic theory of reasoning of Koralus & Mascarenhas (2013) offers an account of this and related illusory inferences that proposes that a question-answer dynamic is at the core of these fallacies. For ease of exposition, let us consider the logical structure behind the example in (1).

(2) \((a \land b) \lor c\)
   \(a\)
   **Fallacious conclusion**: \(b\)
‘Might’ as a generator of alternatives

Following Hamblin semantics (Kratzer & Shimoyama 2002; Alonso-Ovalle 2006) and inquisitive semantics (Groenendijk 2008; Mascarenhas 2009), the erotetic theory of reasoning takes the disjunction in the first premise of (2) to raise an issue, putting forth two alternatives: $a \land b$ and $c$. The reasoner is now effectively entertaining a question, and she will seek to find the most expedient way of answering it or dispelling it. As it turns out, the second premise offers a hint of an answer: the second premise $a$ is related to the first alternative $a \land b$ rather than the second $c$, and the reasoner rushes to pick an answer: the right alternative from the first premise is $a \land b$, whence $b$ follows by a mental-models analog of conjunction elimination.

1.3 Alternatives in language beyond disjunction

If the attractiveness of fallacious schemata as in (2) is to be explained in terms of the presence of alternatives in the interpretation of the first (disjunctive) premise, then we expect other linguistic items that have been argued to raise alternatives of a similar kind to produce similar illusory inferences.

Both in the Hamblin semantics (Kratzer & Shimoyama 2002) and the inquisitive semantics (Ciardelli 2009) traditions, indefinites are akin to $wh$-questions. An expression such as “some pilot” is assigned roughly the same meaning as “which pilot,” namely the set of all pilots. Accordingly, a sentence like “Some pilot writes poems” is analyzed as the set of all propositions of the shape “$x$ writes poems,” for $x$ a pilot. In other words, the meaning of the question “Which pilot writes poems?” Under this analysis, the erotetic theory predicts that the example below should give rise to an illusory inference.

---

1 Strictly speaking, the Hamblin semantics tradition postulates an existential quantifier over alternatives at the very top of any logical form, ensuring that classical truth conditions are delivered instead of sets of alternatives. It is inquisitive semantics that dispenses with this final step. With that said, we think it is reasonable to slightly adjust standard Hamblin semantics and do without existential closure at the top, maintaining existential closure for embedded clauses.

2 This informal exposition of the erotetic theory of reasoning suffices for the purposes of this article, but it is crucial to note that the theory is given in a fully explicit form by Koralus & Mascarenhas (2013). They give a full regimentation of mental model theory in terms of truth-maker semantics (van Fraassen 1969; Fine 2012) with an inquisitive / Hamblin semantics for disjunction, and they define a small set of dynamic operations on mental models. The central operation is Q(uestion)-Update, which checks whether the workspace of reasoning contains a mental model with more than one alternative (a question) and attempts to answer it with the information in the mental-model interpretation of the premise being integrated. In the original (2013) formulation of the theory, this operation simply checked for overlap: the alternatives in the question that had the most overlap in content with the information in the answer were kept, while all others dropped from attention. More recent experimental data have shown that overlap is too strong a requirement (Sablé-Meyer & Mascarenhas 2019). The novel data in the present article do not require these more sophisticated theories, so we use the original formulation of the theory.

---
Some pilot writes poems.
John is a pilot.

**Fallacious conclusion:** John writes poems.

Assuming a finite domain, we can think of the meaning of the first premise of (3) as a large disjunction of conjunctions. This is given schematically in (4), for $P$ a predicate for “pilot” and $W$ a predicate for “writes poems.” Notice how structurally we have a generalized case of the illusory inference with disjunction in (2).

(4)  $(Pa \land Wa) \lor (Pb \land Wb) \lor \cdots \lor (Pz \land Wz)$
$Pa$

**Fallacious conclusion:** $Wa$

**Mascarenhas & Koralus (2017)** found that inferences like (3) were in fact attractive, with acceptance rates around 35%, significantly above the baseline for mistakes established by invalid controls without alternative-generating elements. This demonstrates the existence of the predicted illusory inferences with alternative-generating linguistic operators besides disjunction.

### 1.4 Reasoning vs. interpretation

Illusory inferences from *disjunction* as in (1) have acceptance rates around 85%, while illusory inferences with *indefinites* as in (3) are accepted around 35% of the time. Why the different acceptance rates, if both disjunctions and indefinites produce alternatives according to our best semantic theories?

**Mascarenhas (2014)** showed that the disjunction inferences are amenable to a pragmatic account. As an alternative to mental-models accounts, we can predict the observed conclusion assuming an entirely classical reasoning module, acting on pragmatically strengthened meanings of the premises. The computations are far from easy, but on most modern theories of scalar implicature (e.g. Sauerland 2004; Spector 2007) the first premise of the illusory inference from disjunction in (5a) is predicted to be interpreted as in (5b), a fact Spector (2007) had already observed outside the context of reasoning problems.

(5)  a. $(a \land b) \lor c$
    b. $(a \land b \land \neg c) \lor (c \land \neg a \land \neg b)$

Assuming (5a) is interpreted as in (5b), and incorporating the second premise $a$, the conclusion $b$ is no fallacy at all. It follows classically from the conjunction of the two premises. Crucially, no absolving implicature is predicted by any theory of scalar implicature in the market for illusory inferences with *indefinites*. **Picat & Mascarenhas (2019)** argue that, rather than being competing accounts, the erotetic
theory and the scalar-implicature account are two possible routes leading to the same inference-making behavior in the case of the illusory inference from \textit{disjunction}. For the indefinites case, only the erotetic theory predicts a fallacy, the scalar-implicature route is blocked, explaining its lower endorsement rate.\textsuperscript{3}

\section{Illusory inferences with ‘might’}

Ciardelli \textit{et al.} (2009) argue that the weak epistemic operator ‘might’ is an alternative-generating operator, putting forth a single alternative and drawing attention to it. Ciardelli \textit{et al.} (2009) give an inquisitive semantics for ‘might’, where the meaning of a sentence $\text{might}(p \land q)$ is $(p \land q) \lor \top$. In inquisitive semantics, this corresponds to the question $\{p \land q, \top\}$. Crucially, the sentence as a whole is not a tautology: while it is informationally idle (that is an update with this sentence will exclude no possibilities from any common ground), it contains two alternatives, and is thus distinct from the interpretation of a sentence $\top$.

This inquisitive semantics predicts two alternatives for ‘might’-sentences, but one of those alternatives is the trivial alternative, to which it makes little or no sense to draw attention. Effectively then, a ‘might’-sentence in this semantics offers \textit{one} alternative to focus on, in the example under consideration the $p \land q$ alternative.

If indeed ‘might’ raises an alternative in the sense of inquisitive semantics and Hamblin semantics, then ‘might’ should produce illusory inferences in the right configuration. Schematically, we expect the pattern in (6) to be attractive.

\begin{equation}
\text{might}(a \land b) \\
\quad a \\
\text{Fallacious conclusion: } b
\end{equation}

\subsection{Design}

We tested this hypothesis experimentally. Our target reasoning problems had the structure in (6), instantiated in (7).

\begin{equation}
\text{Miranda might play the piano and be afraid of spiders.} \\
\text{Miranda plays the piano.} \\
\text{Fallacious conclusion: Miranda is afraid of spiders.}
\end{equation}

Earlier experimental investigations on illusory inferences from disjunctions and indefinites found a small but significant effect of premise order: fewer fallacies

\textsuperscript{3} Picat \& Mascarenhas (2019) give corroborating experimental evidence for this proposal, showing that the disjunction case is affected by cognitive load, reducing pragmatic interpretations and decreasing the rate of fallacious responses, while the indefinite case is unaffected by this manipulation.
were committed when the premises were presented in reverse order, in line with the erotetic theory’s account in terms of question-answer dynamics (Mascarenhas & Koralus 2017; Koralus & Mascarenhas 2018). Accordingly, we tested inferences like (7) in reverse order as well.

Fallacious conclusions in response to an inference like (7) could be due to two confounds we controlled for. Firstly, it could be that the first premise alone prompts a fallacy, perhaps by raising the probability of the conjunction sufficiently in the judgment of participants. This would still be of interest, but it would not constitute an illusory inference of the kind we find with disjunctions and indefinites. Secondly, we wanted to make sure that any mistakes we found were due to a bona fide question-answer dynamic, as predicted by the erotetic theory. Consequently, we tested as controls both the first premise alone, and what we call a “flat” version of the inference with the same context-change potential, exemplified in (8).

(8) Miranda plays the piano and might be afraid of spiders.

**Fallacious conclusion:** Miranda is afraid of spiders.

It is useful to contrast the reversed version of the inference with this flat version, schematically:

(9) a. **REVERSED TARGET**
   
   \[
   a \\
   \text{might}(a \land b)
   \]

   **Fallacious conclusion:** \[b\]

b. **FLAT**

\[
\neg a \land \text{might}(b)
\]

**Fallacious conclusion:** \[b\]

In the context of the first premise \(a\) of (9a), an utterance of \(\text{might}(a \land b)\) should be contextually equivalent to \(\text{might}(b)\). The idea behind the flat version in (9b) is to exploit this contextual equivalence: the two discourses in (9) should yield identical results, unless something like the question-answer dynamic postulated by the erotetic theory of reasoning is at play.

Summing up, we tested the following four conditions in a between-subjects design, due to the high degree of similarity between the stimuli in the conditions.

(10) a. **CANONICAL**

\[
\text{might}(a \land b), a \vdash b
\]

b. **P1**

\[
\text{might}(a \land b) \vdash b
\]

c. **FLAT**

\[
a \land \text{might}(b) \vdash b
\]

d. **REVERSED**

\[
a, \text{might}(a \land b) \vdash b
\]

We further tested valid and invalid controls based on modus ponens, as schematized in (11).
`Might` as a generator of alternatives

\[(11)\]
\begin{align*}
\text{a. VALID MP} & \quad \text{if } a \text{ then } b, a \vdash b \\
\text{b. INVALID MP} & \quad \text{if } a \text{ then } b, c \vdash b
\end{align*}

We recruited 210 subjects on Amazon Mechanical Turk, 66% were female and their mean age was 36 (ranging from 18 to 74, \(\sigma = 11.4\)). Participants were assigned to one of the four target conditions. They each solved 14 reasoning problems, 8 targets and the rest controls. In each case, participants were given premises and a proposed conclusion, and asked whether the conclusion followed from the premises. We gave participants two examples in the instructions to the experiment, exemplifying one valid and one invalid inference, with materials and connectives that did not occur elsewhere in the experiment.

We predicted that

i. canonical and reversed targets should be accepted significantly more than the baseline for mistakes established by invalid controls, diagnosing a fallacy;

ii. the acceptance of canonical and reversed targets should depend on the presence of the second premise, so that P1 targets should be lower than canonical and reversed targets;

iii. flat targets should be much lower than canonical and reversed targets, and indeed indistinguishable from the baseline for mistakes established by invalid controls;

iv. canonical targets should be somewhat more attractive than reversed targets, instantiating the order effects observed for other illusory inferences.

2.2 Results

We excluded from the analysis participants who reported having taken graduate-level courses in natural language semantics or pragmatics, and participants who reported using notes during the experiment. In total, we excluded 64 participants. We analyzed the remaining data using Wilcoxon signed-rank tests on the proportion of yes-answers given to each type of inference and control. The details of the tests are given in Table 1, while Figure 1 summarizes the results.

2.3 Discussion

Most of our predictions were borne out:

i. Canonical and reversed targets were significantly more accepted than mistakes were made on no-controls. This is diagnostic of a reasoning fallacy.
Table 1  Details of the statistical analysis; all comparisons were analyzed with Wilcoxon signed-rank tests; for comparison (i) we performed a paired Wilcoxon signed-rank test (within subjects)

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>$V = 805$</td>
<td>$p &lt; 0.01$</td>
</tr>
<tr>
<td>(ii)</td>
<td>$W = 1485$</td>
<td>$p &lt; 0.05$</td>
</tr>
<tr>
<td>(iii)</td>
<td>$W = 1683.5$</td>
<td>$p &lt; 0.001$</td>
</tr>
<tr>
<td>(iv)</td>
<td>$W = 533.5$</td>
<td>$p &gt; 0.8$</td>
</tr>
</tbody>
</table>

Figure 1  Results for the relevant conditions in the experiment; error bars indicate the standard error of the mean

ii. Canonical and reversed targets were significantly more accepted than P1 targets. Thus, the fallacy cannot be explained by the first premise alone.

iii. Canonical and reversed targets were significantly more accepted than flat targets. This is in line with an erotetic account of the fallacy.

iv. However, we detected no difference between canonical and reversed targets, suggesting no order effect. We can understand this null result by appealing to effect size. Experiments showing the existence of order effects for fallacies with similar acceptance rates required very large numbers of participants due to the small effect size observed (Mascarenhas & Koralus 2017). The current experiment was very plausibly underpowered to answer this particular question.
‘Might’ as a generator of alternatives

3 General discussion

We’ve concluded that illusory inferences with ‘might’ exist, as predicted by accounts of ‘might’ as an inquisitive or alternative-generating operator. These cannot be explained away as resulting from the first premise alone, and they seem to require a dynamic operation as evidenced by the absence of fallacies for flat targets.

What do the results of our experimental investigation entail for existing analyses of the semantics of ‘might’, and existing theories of how alternatives figure into human reasoning? We turn our attention to four relatively broad possibilities, which we believe map out a considerable portion of the relevant conceptual space.

3.1 Inquisitive semantics and the erotetic theory of reasoning

As discussed in the introduction to this article, Ciardelli et al. (2009) offer an inquisitive semantics for ‘might’ where

\[ \text{might}(\varphi) \equiv \varphi \lor \top. \]

Armed with this semantics, we need only import the propositional erotetic account of reasoning of Koralus & Mascarenhas (2013) to derive the observed fallacious behavior as a special case of the original illusory inference from disjunction in (1).

3.2 Scalar Implicatures and interpretation-based “fallacies”

Illusory inferences from ‘might’ have comparable acceptance rates to those with indefinites, rather than disjunctions. Following Picat & Mascarenhas (2019), we expect these inferences to have only one source, erotetic reasoning. That is, there should be no scalar implicature of either premise that derives the “fallacy” as a result of valid reasoning. Indeed, in order to validate the observed conclusion, we would need for the first premise of the inference to have the implicature in (12).

\[ \Diamond (a \land b) \land \neg \Diamond (a \land \neg b) \]

i.e. \[ \Diamond (a \land b) \land \Box (a \to b) \]

We know of no theory of pragmatics that would predict such an implicature for the first premise of our inferences with ‘might’, confirming our expectations.

3.3 Relational semantics

Assuming the relevant ordering between possible worlds is well founded (the limit assumption), Kratzer’s (1991) semantics for modality predicts that a sentence \textit{might}(\varphi)
Mascarenhas and Picat

will be true just in case there is a \( \varphi \)-world among the best-ranked worlds. The existential quantifier over possible worlds in these truth conditions can be given an inquisitive semantics just like that of indefinite expressions. Along with a few plausible assumptions, there is a strategy for accounting for our data in a relational semantics such as Kratzer’s.

(13) a. **Assertion as truth in the actual world**

When asserting a proposition \( \varphi \), a speaker communicates (their belief) that \( \varphi \) holds in the actual world.

b. **Reflexivity**

The epistemic modal base is reflexive, that is \( w \in f(w) \).

c. **Inquisitiveness of existentials and erotetics**

The existential quantifier that occurs in the truth conditions of a sentence with ‘might’ is inquisitive. Human reasoning is erotetic, and in particular inquisitive existential quantifiers raise questions that are resolved along the lines sketched in section 1.3 for illusory inferences with indefinites.

Under these assumptions, our first premise \( \text{might}(a \land b) \) provides some information about the existence of a possible world, and draws attention to the question “which is this best-ranked \( a \land b \)-world we’re discussing?” Asserting the second premise then states that the actual world is an \( a \)-world. Now erotetic mechanisms kick in: the actual world is an \( a \) world, and this points in the direction of answering the question at hand with “the actual world is the best-ranked \( a \land b \)-world in question.” From here \( b \) follows immediately.

### 3.4 Probabilistic semantics

Lassiter (2016) proposes a probabilistic semantics for ‘might’, where

\[
\text{might}(\varphi) \equiv P(\varphi) > \theta ,
\]

for \( P \) a probability function and \( \theta \) a contextually determined standard. To assess how this semantics handles our inferential data, we need to formulate a theory of reasoning with probabilities. One standard notion of validity in probabilistic approaches to the psychology reasoning is \( p \)-validity (Oaksford & Chater 2007; Johnson-Laird, Khemlani & Goodwin 2015). An inference is \( p \)-valid just in case the conclusion is no less probable than its premises. That is, for \( \varphi \) a conclusion and \( \Gamma \) a set of premises,

\[
P(\varphi) \geq P(\Gamma) .
\]

Under this operationalization of probabilistic reasoning, we see no principled obstacle to a probabilistic semantics for ‘might’ incorporating our findings. However,
‘Might’ as a generator of alternatives

the notion of $p$-validity has mostly been used to great effect in accounts of conditional reasoning, but it is otherwise a problematic notion. The issue deserves a detailed discussion that is outside the scope of this article. But we can illustrate the central oddity with an example: take a fair lottery with 100 tickets, numbered from 1 to 100. From “the winning ticket is 3” it follows by $p$-validity that “the winning ticket is 4 or 5.” Relatedly, any proposition either entails its negation, is entailed by it, or both. We submit that these predictions are almost certainly on the wrong track.

Alternatively, one can consider a more conservative notion of probabilistic reasoning using conditional probabilities. For $\varphi$ a conclusion, $\Gamma$ a set of premises, and $\tau$ a contextually determined standard, we could say that $\varphi$ follows from $\Gamma$ just in case

$$P(\varphi | \Gamma) > \tau.$$  

As before, a detailed discussion cannot quite be dispensed with. But we note that the problems raised above for $p$-validity needn’t arise for this notion of probabilistic reasoning. If the standard $\tau$ is reasonable (that is high) enough, the problematic cases will not be valid. For the lottery case for example, conditionalizing on the winning ticket being number 3 will in fact reduce the probability of the winning ticket being 4 or 5 to zero.

But this move comes with its own open questions. In (14a) we spell out the predicted acceptance conditions for canonical and reversed targets, and in (14b) the predicted acceptance conditions for flat targets.

(14) a. CANONICAL / REVERSED

$$P(b|a \& P(a \land b) > \theta) > \tau$$

b. FLAT

$$P(b|a \& P(b) > \theta) > \tau$$

The issue is whether the two conditional probabilities can be distinguished. One would need to spell out and look at the details of a semantics for a probability language with second order probabilities. But at first glance, it is reasonable to wonder whether, upon conjunction with $a$, the probability of $a \land b$ should reduce to that of $b$ alone. If so, then the two sets of premises might well be equivalent, which would make them indistinguishable, going against the results we report in this article. The question is open as far as we can see, and deserves attention.

4 Conclusions

We have shown that ‘might’ induces illusory inferences comparable to those produced by indefinites. This constitutes the first experimental argument in favor of a theory of ‘might’ along the lines of what’s proposed by Ciardelli et al. (2009). In a
nutshell, ‘might’ is an alternative-generating element, drawing attention to a single possibility. In ongoing work, we are investigating other weak modals with the same methods, as well as other linguistic elements that have been argued on theoretical grounds to be generators of alternatives.

One speculative but promising conjecture emerges from this work. It is clear that linguistic expertise is essential in the study of the psychology of human reasoning. Virtually the totality of work in this field uses linguistically presented reasoning and decision problems to diagnose failures of reasoning, so a thorough understanding of interpretive processes is indispensable. But the methods of the psychology of reasoning can be illuminating to the field of natural language semantics as well. If our theories and our interpretation of these experimental results are on the right track, then illusory inferences of the sort discussed here can serve as novel diagnostic tools for the presence of alternative-generating elements in semantics.

References

‘Might’ as a generator of alternatives


Salvador Mascarenhas
Ecole Normale Supérieure
Department of Cognitive Studies
29 rue d’Ulm
75005 Paris, France
salvador.mascarenhas@ens.fr

Léo Picat
Ecole Normale Supérieure
Department of Cognitive Studies
29 rue d’Ulm
75005 Paris, France
picatleo@hotmail.fr

561