

Illusory inferences with quantifiers

(preprint of article accepted for publication by *Thinking and Reasoning*)

Salvador Mascarenhas and Philipp Koralus*

Laboratory for the Philosophy and Psychology of Rationality and Decision
Faculty of Philosophy, University of Oxford

February 24, 2016

Abstract

The psychological study of reasoning with quantifiers has predominantly focused on inference patterns studied by Aristotle about two millennia ago. Modern logic has shown a wealth of inference patterns involving quantifiers that are far beyond the expressive power of Aristotelian syllogisms, and whose psychology should be explored. We bring to light a novel class of fallacious inference patterns, some of which are so attractive that they are tantamount to cognitive illusions. In tandem with recent insights from linguistics that quantifiers like “some” are treated as *wh*-questions, these illusory inferences are predicted by the erotetic theory of reasoning, which postulates that a process akin to question asking and answering is behind human inference making.

Keywords: reasoning; quantifiers; syllogisms; mental models; questions

1 Introduction

Since the birth of modern logic in the late 19th century and the pioneering work of Gottlob Frege and Bertrand Russell, both propositional and quantified inference patterns have been studied in great detail mathematically (Frege, 1879; Whitehead and Russell, 1910). However, the psychological study of reasoning with quantifiers has predominantly placed its focus about two millennia earlier: Aristotelian syllogisms (Ross, 2000; Chater and Oaksford, 1999; Wetherick and Gilhooly, 1995; Ford, 1995; Geurts, 2003; Guyote and Sternberg, 1981; Bucciarelli and Johnson-Laird, 1999).

In this article, we present a novel paradigm of attractive fallacious inferences with quantifiers that do not constitute syllogisms. The inference in (1), with two universal quantifiers “every king” and “every queen,” was accepted by an overwhelming majority of subjects in an experiment we report on in this article.

* The work reported in this article was supported by grants from the Laces Trust and the John Fell Fund.

- (1) P₁: Mary has met every king or every queen of Europe.
P₂: Mary has met the king of Spain.
Concl.: *Does it follow that Mary has met the king of Belgium?*

But the novel inference patterns are not restricted to universal quantifiers. The inference in (2), with the indefinite quantifier “some pilot,” was accepted by a sizable portion of subjects, albeit a smaller proportion than accepted the inference with universals in (1).

- (2) P₁: Some pilot writes poems.
P₂: John is a pilot.
Concl.: *Does it follow that John writes poems?*

The patterns in (1) and (2) bear interesting similarities to illusory inferences with disjunction first introduced by Walsh and Johnson-Laird (2004). The parallelism builds on the fact that an indefinite quantifier as in (2) can be seen mathematically as a generalized disjunction (A is a pilot who writes poems or B is a pilot who writes poems or . . .). The following is a representative example of these disjunction inferences.

- (3) P₁: Either Jane is kneeling by the fire and she is looking at the TV or otherwise Mark is standing at the window and he is peering into the garden.
P₂: Jane is kneeling by the fire.
Concl.: Jane is looking at the TV.

In a study by Walsh and Johnson-Laird (2004), over 80% of subjects considered that the conclusion in (3) follows from the premises. But according to classical logic this is not the case, for it is compatible with the premises but not with the conclusion that Jane be kneeling and not looking, while Mark is both standing and peering. Notice moreover that the conclusion in (3) is fallacious even if we interpret ‘or’ as an exclusive disjunction. That is, the conclusion does not follow even if “A or B” means “A or B but not both.” This is because, schematically, the first premise says “[A and B] or [C and D].” If we interpret ‘or’ as exclusive, we get “[A and B and not [C and D]] or [C and D and not [A and B]].” But incorporating the second premise “A” into this exclusive disjunction still does not validate the conclusion “B.”

Inferences as in (3) were accounted for by Walsh and Johnson-Laird (2004) in terms of the mental models theory of reasoning (Johnson-Laird, 1983). We observe that, while the mental models theory approach to reasoning with quantifiers can in principle be applied to non-syllogistic data with first order quantifiers as presented in this article, the extant experimental literature has not done so, focusing instead of Aristotelian syllogisms. This should be cause for some concern, for syllogisms represent a minute portion of the full expressive power of a language with quantifiers such as English or first-order logic. Consequently, we run the risk of tailoring our theories of reasoning to a subset of the phenomenon that is not representative enough. The novel data we report on constitute first steps toward a necessary widening of our empirical domain.

The erotetic theory of reasoning of Koralus and Mascarenhas (2013), a variant of mental model theory, explains the inference in (3) and predicts that indefinite quantifiers

should give rise to similar fallacies. We begin with a brief introduction to the erotetic theory of reasoning. As far as we can see, the erotetic theory is the only theory that makes immediate predictions for the quantified cases we consider in this article. In section 2 we present and discuss three experiments designed to check the predictions of the erotetic theory for these novel quantified cases. Section 3 concludes with a more detailed overview of our account and a comparison with possible alternative accounts.

1.1 The erotetic theory of reasoning

Though our concern in this article is to present experiments establishing a novel class of attractive fallacious inferences with quantifiers, it is important to highlight the sense in which our new data with quantifiers are related to the more familiar illusory inferences from disjunction. On our erotetic (i.e., question-based) account, to be discussed in more detail in this article's general discussion section (3.1), reasoners are primarily engaged in a question asking and answering task. But the sources of questions taken onboard by reasoners are often non-obvious. In particular, a disjunctive sentence like " A or B " on our account will not only provide information but also pose a question: "which of A and B is indeed the case?" Consider the example in (3), from the mental models literature on illusory inferences (Walsh and Johnson-Laird, 2004). On the erotetic view of the semantics of "or," the first premise poses the question "are we in a Jane kneeling and looking situation or in a Mark standing and peering situation?" The moment reasoners are attending to a question, they try their best to resolve it swiftly, considering every later premise as much as possible as a strong answer. Premise 2 of (3) provides an incomplete answer: "we are in a kneeling situation." At this point, the reasoner observes that the putative answer has more in common with the first possible answer to the question at hand than the second. She concludes: "we are in a Jane kneeling and looking situation," whence the fallacious conclusion that "Jane is looking at the TV" follows immediately.

In section 3.1 we give a more detailed overview of the theory. The formal system we define elsewhere (Koralus and Mascarenhas, 2013) accounts for a host of existing empirical data points from the literature on reasoning with propositional connectives, including the original illusory inferences from disjunction, and it allows one to calculate predictions for novel inference patterns of arbitrary complexity. The theory given in that article has also yielded insights on monothematic delusional thinking (Parrott and Koralus, 2015). Although Koralus and Mascarenhas (2013) give an account of propositional reasoning and not yet of quantified reasoning, there are substantive predictions that the theory in its present form already makes for inferences involving indefinite quantifiers.

1.2 Predictions for quantifiers

Several sources in linguistic semantics have championed the view that indefinite expressions share with disjunctions and interrogatives the ability to raise a question, in particular, alternative/Hamblin semantics (Kratzer and Shimoyama, 2002) and inquisitive semantics (Ciardelli, 2009; Mascarenhas, 2011). In other words, there are good independent reasons to think that (4a) shares with (4b) the ability to raise a question.

- (4) a. Some pilot writes poems.
- b. Which pilot writes poems?

With the erotetic theory of reasoning, this makes the following prediction: it ought to be possible to follow (4a) with a categorical (i.e. non-disjunctive and non-indefinite) premise that partially overlaps with exactly one of the possible answers to (4b), giving rise to a fallacious conclusion. The following is a good candidate.

- (5) Some pilot writes poems.
John is a pilot.
Does it follow that John writes poems?

On this view, premise 1 of (5) raises the question “who is a pilot who writes poems?” Following standard approaches to the semantics of questions, this has as possible answers propositions of the form “So-and-so is a pilot and writes poems,” one for each so-and-so in the domain of quantification.

Then, premise 2 is interpreted as the strongest possible answer to this question. Premise 2 overlaps with only one of the answers to the question, namely “John is a pilot who writes poems,” assuming, as is natural, that “John” is within the domain of quantification. Following the erotetic account, reasoners should take the question to have been answered: “John is a pilot and writes poems,” whence the fallacious conclusion that John writes poems follows immediately. Crucially, the pattern exemplified in (5) does not constitute an Aristotelian syllogism, and thus has escaped the consideration of most psychological theories of reasoning with quantifiers.

2 Illusory inferences with quantifiers

We tested for non-syllogistic, indefinite versions of illusory inferences from disjunction with two experiments.

2.1 Pilot experiment — familiar terms

In a pilot experiment, we tested the two patterns in (6):

- (6) a. Some pilot writes poems.
John is a pilot.
Does it follow that John writes poems?
- b. Some student smokes.
John is a student.
Does it follow that John smokes?

In the aggregate, we found that over 30% of subjects made the fallacious conclusion, significantly more than our invalid controls (Wilcoxon matched-pairs test $p < .05$). However, a by-item analysis showed that the effect was only significant for (6a).

2.2 Experiment 1 — unfamiliar terms

The pilot just described showed a suspicious amount of variance in the targets: some of our target stimuli displayed a rate of acceptance significantly higher than our invalid controls, some did not. In experiment 1 we addressed this issue. We hypothesized that subjects' prior expectations about what is typical for the topical individuals of our materials were responsible for the variance. For example, we suspected that the easy availability of known examples made the corresponding illusory inference seem less attractive. On the erotetic theory (Koralus and Mascarenhas, 2013), it is expected that the ease with which reasoners can conjure up alternative representations based on their background knowledge should correlate inversely with the attractiveness of the relevant fallacious inferences. We decided to factor out entirely subjects' expectations about the topical individuals, by using completely unfamiliar properties in both restrictor and nuclear-scope positions in our materials.

2.2.1 Design and participants

We recruited 976 workers from the mTurk crowdsourcing platform to solve twelve reasoning problems. Forty-eight percent of our participants were female, and the sample's average age was 32 years (18-76 years, $\sigma = 10.8$). Statements in each reasoning problem were about biological organisms and properties (see materials section below). Of the twelve problems subjects saw, six constituted variations on our target illusory inferences with indefinites, and six were control inferences, three valid and three invalid. Our target stimuli were further divided between a group in canonical order — premise with an indefinite expression first — and in reversed order. Valid controls were inferences that the literature has suggested reasoners have little to no difficulty with. Our invalid controls were invalid inferences involving the indefinite quantifier “some” that did not conform to the question-answer configuration that the erotetic theory predicts should induce reasoners into fallacious reasoning. Each subject solved all twelve problems, serving as his or her own control, in a randomized order. Subjects were asked not to make use of search engines while answering the questionnaire.

2.2.2 Materials and methods

To control for the role of subjects' prior expectations, our materials used expressions from biology to form grammatical sentences of English with the required structure. We concocted the examples ourselves with no regard for the truth of the sentences or coherence of the statements as statements about biology. This was a legitimate strategy given that we presumed no knowledge of the subject matter. In (7) we give two representative content variations we used for canonical targets, and in (8) two examples of invalid controls. Notice that the invalid controls in (8) do not conform to the question-answer configuration seen in the targets in (7): in the invalid controls, the categorical premise, being negative, does not contribute an answer to the question raised by the indefinite.

- (7) a. Some firmicute produces endospores.
Clostridium is a firmicute.
Does it follow that Clostridium produces endospores?
- b. Some thermotogum stains gram-negative.
Maritima is a thermotogum.
Does it follow Maritima stains gram-negative?
- (8) a. Some dictyoglomus is thermophobic.
Turgidum is not a dictyoglomus.
Does it follow that Turgidum is thermophobic?
- b. Pectinis is not a spirochetes.
Some spirochetes undergoes transverse fission.
Does it follow that Pectinis undergoes transverse fission?

Our materials also included two stimuli with mixed quantifiers, involving both an indefinite and a universal quantifier, as exemplified in (9)

- (9) Some bacillus metabolizes every sulfide in its environment.
Heliobacillus metabolizes hydrogen sulfide.
Does it follow that Heliobacillus metabolizes dimethyl sulfide?

Participants had to actively select their answer (“yes,” “no,” or “choose not to answer”) from a dropdown list and click on a “Next” button to proceed. The last page of the questionnaire collected basic demographic information: age group, gender, background in logic, background in biology, use of notes and search engines.

<i>Type of stimulus</i>	<i>N mistakes (%)</i>
Target a — canonical order	374 (38.3%)
Target b — canonical order	384 (39.3%)
Target c — reversed order	300 (30.7%)
Target d — reversed order	328 (33.6%)
Mixed quantifiers a	95 (9.7%)
Mixed quantifiers b	218 (22.3%)
Valid controls — canonical order	70 (7.2%)
Valid control — reversed order	75 (7.7%)
Invalid control a — canonical order	110 (11.2%)
Invalid control b — canonical order	104 (10.7%)
Invalid controls — reversed order	120 (12.3%)

Table 1: Experiment 1 main results

2.2.3 Results

Table 1 above summarizes our findings. We found that 40% of subjects committed the predicted fallacy. This was a significantly higher rate of reasoning error than in our invalid controls, which contained indefinite expressions but did not instantiate a question-answer configuration (Wilcoxon matched-pairs test $V = 89107.5$, $p < .01$). Unlike in our pilot, in experiment 1 we found no effect of item content. We found no significant effect of subjects’ background in biology ($\chi^2 < 8.6$, $p > .19$), but there was a significant trend

where subjects with more background in logic, philosophy, or linguistics endorsed the fallacy less frequently ($\chi^2 > 15.5$, $p < .02$). Less than 10% of participants admitted to having resorted to notes or diagrams while answering the questionnaire, and only 5% admitted to using search engines. Our two stimuli with mixed quantifiers displayed lower acceptance rates and wide variance. We hypothesize that the added complexity of dealing with two quantifiers, as well as the introduction of scopal ambiguity, made these materials especially difficult.

We also found a significant order effect on the premises: when the order of the premises in stimuli as in (7) was inverted, we found significantly fewer fallacies for target stimuli. The drop in acceptance was just under 10%, and it was statistically very significant (Wilcoxon matched-pairs test $V = 28836.5$, $p < .01$). This order effect was not significant in either our valid controls ($p > .61$) or our invalid controls ($.17 < p < .41$).

2.2.4 Discussion

These results fit with the prediction of the erotetic theory: indefinites give rise to fallacies akin to those of disjunction. These fallacious inferences are novel, and do not constitute syllogisms. They are straightforwardly accounted for by the erotetic theory of reasoning, naturally extended to a first order language.

The order effect we found is also naturally explained by the erotetic theory. If the fallacies in (7) are a product of the particular dynamic of question asking and answering, we expect that fewer people would make the fallacious inference when they are given first the answer and then the question. The fact that our valid and invalid controls, which did not conform to the question-answer schema, did not display a significant effect of premise order offers corroborating evidence for an erotetic account. We should add that an erotetic account does not predict that it should be impossible to get a fallacy for our reverse-order targets. Some reasoners might find that nothing novel can be concluded from the problem as given, and then attempt it again by reversing the order of the premises for themselves. Crucially, other theories that explain reasoning data germane to ours fail to explain this order effect. In particular, standard mental model theory lacks a procedure for combining these premises that is asymmetric in the right way.

2.3 Experiment 2 — universal quantifiers

Our pilot suggested, and experiment 1 confirmed, that there exist robust fallacious inferences with quantifiers beyond syllogisms. We suspected that the phenomenon was not restricted to indefinites. In experiment 2, we made a different “translation” of the standard illusory inference in (3) into the realm of quantification, building on an idea from Mascarenhas (2014). Instead of using indefinites to do the work of disjunction, we used universal quantifiers where (3) has conjunctions. We used five target items, of which (10a) and (10b) below are representative examples. Control problems had the same properties as in experiment 1 discussed above and are exemplified in (11). For this experiment, we did not look for an order effect of premises, and thus our materials only included targets in canonical order (universal premise first). Also as with the patterns with indefinites, notice that the inferences in (10) do not constitute Aristotelian syllogisms.

- (10) a. Every boy or every girl is coming to the party. John is coming to the party.
Does it follow that Bill is coming to the party?
- b. Mary has met every king or every queen of Europe. Mary has met the king of Spain.
Does it follow that Mary has met the king of Belgium?
- (11) a. Some teachers are about to retire.
Eva is not a teacher.
Does it follow that Eva is about to retire?
- b. Jane is not German.
Some Germans follow soccer.
Does it follow that Jane follows soccer?

The method was entirely parallel to the one used in experiment 1, except that we recruited a new and smaller set of participants ($N = 203$, 34% female, mean age 33 years). As predicted, we found acceptance rates comparable to those reported by Walsh and Johnson-Laird (2004) for the propositional counterpart, at around 82%. This was significantly more than the frequency of mistakes in valid (under 5%) or invalid (10%) controls (Wilcoxon matched-pairs test $p < .0001$). Given that the materials in (10) involve disjunction and a very high acceptance rate, they are close parallels of the original illusory inferences from disjunction studied in the mental models framework. This result demonstrates that the range of illusory inferences is not limited to propositional connectives or indefinite expressions.

3 General discussion

3.1 The erotetic theory of reasoning

At the core of the erotetic theory is the erotetic principle, as follows.

- (12) The erotetic principle
- Part I** — Our natural capacity for reasoning proceeds by treating successive premises as questions and maximally strong answers to them.
- Part II** — Systematically asking a certain type of question as we interpret each new premise allows us to reason in a classically valid way.

The erotetic theory of reasoning is a descriptive and explanatory theory of human inference-making behavior. It addresses the problem of failure of human reasoning—how and why do humans systematically err when answering many ostensibly elementary logical problems?—by proposing that, in many representative instances of failures of reasoning, those failures are due to the fact that certain superficially declarative premises pose questions, to which humans respond by looking for answers in other premises, guided by a desire to resolve open questions as soon as possible. This is Part I of the erotetic principle in (12).

Our account is also concerned with the problem of success in human reasoning: if we are prone to such a broad array of fallacious inference-making behavior, how is logically sound reasoning possible at all? Are there cognitive mechanisms that we can reliably

deploy in order to guarantee normatively sanctioned reasoning? Part II of the erotetic principle addresses the issue of success. Resorting to a particular but natural strategy of asking further questions at key points during reasoning, humans can match logically sound inference making, albeit at a processing cost.

The account is fully developed for reasoning with propositional connectives: Koralus and Mascarenhas (2013) give a formal system that implements the erotetic desiderata in (12). Concerns of space prevent us from presenting that formal system in this article, but in what follows we summarize the key elements of the account of both failures and successes of reasoning.

3.1.1 The problem of failure

Consider again the illusory inferences from disjunction of Walsh and Johnson-Laird (2004) in (3) above. Our account imports from a current of linguistic theories of the semantics of disjunction the insight that some superficially declarative sentences share with questions the property of raising an issue (Groenendijk, 2008; Mascarenhas, 2009). Note that on this view taking on board a question or issue as a result of hearing a premise statement is not just a pragmatic enrichment or conversational phenomenon, but straightforwardly what the linguistic meaning of the statement delivers to our conceptual/intentional systems. Correspondingly, on the erotetic theory, the question-answer character of reasoning is a property of human thinking, not just narrowly of utterance interpretation. Disjunctions are the paradigmatic case. Accordingly, on the erotetic theory, the first premise of (3) does two things: on the one hand it informs the hearer/reader that at least one of the two situations described must be the case; but in addition to that it raises the issue of which of the two situations is in fact the case. In our view, to understand the meaning of premise 1 of (3) involves taking on board the question: “are we in a Jane kneeling and looking situation or in a Mark standing and peering situation?” We take it that humans dislike entertaining unanswered questions, introducing strong pressure toward resolving the question as swiftly as possible. Thus, the erotetic theory attempts to interpret new premises, in the context of an earlier question, as the strongest possible answer to that question. This desire to resolve open questions is what motivates the general phenomenon of jumping to conclusions.

The erotetic theory is a dynamic account of reasoning: premises are interpreted in the order in which they are given, and at each step reasoners update their mental model discourse, the workspace of reasoning, with the interpretation of each premise, in an operation we call Q-Update (for *Question-Update*). Q-Update is so defined that sentences containing disjunctions (or indefinites, for the quantified case) give rise to two or more alternatives. Whenever the mental model discourse already contains a question, Q-Update attempts to interpret new premises as answers to the question under consideration. Koralus and Mascarenhas (2013) cash out this notion in terms of propositional overlap. We return to the illusory inference in (3) for illustration purposes: interpreted as a strong answer to the question raised by premise 1, premise 2 overlaps with the first possible answer to the question and not the second. In other words, “Jane is kneeling” has more in common with a “Jane kneeling and looking” situation than with a “Mark standing and peering” situation. Via Q-Update, reasoners take premise 2 to answer the question of premise 1 by selecting the first situation: “Jane is kneeling by the fire and looking at the

TV,” thereby dispelling the question and bringing the mental model discourse to a “resting point” as it were.

While Q-Update is responsible for the central job of accounting for jumping to conclusions, the erotetic theory posits alongside it a small number of simple operations on discourses. One such operation is Reduction, which allows a reasoner to go from a representation of A and B to a representation of just A or just B . Applying this operation to the putative answer in our example, reasoners can conclude immediately that “Jane is looking at the TV,” the observed fallacious conclusion.

For completeness, we briefly present three other operations that figure in the theory but are not operative in the fallacies discussed in this article. We take it that humans do not realize immediately (that is “for free”) when they are entertaining contradictions. Accordingly, our theory of mental representations of propositional content allows for alternatives containing contradictions to be represented. We do this in a mathematically sound fashion, resorting to existing tools in the logic literature. We thus define an operation Filter, not part of the default reasoning strategy, that scans the representation in discourse for alternatives containing immediate contradictions and excises them. To allow for hypothetical reasoning and for reasoning with conditionals, we define a Suppose operation that takes on a hypothetical premise (possibly a question), to be discarded with the Depose operation, allowing for conditional conclusions.

3.1.2 The problem of success

The theory is also concerned with the problem of success for reasoning, which is addressed by Part II of the erotetic principle. Humans are not irretrievably lost to the non-normative conclusions that are brought about by a tendency to treat premises as immediate strong answers. In fact, questions play a crucial role in guaranteeing normatively correct reasoning in this account. If reasoners are careful to ask polar questions (i.e. yes-no questions) about each atomic proposition that occurs in the question under consideration before updating with the putative answer supplied by a later premise, it can be formally proved that their reasoning will be classically sound in the technical sense (Koralus and Mascarenhas, 2013). The last operation in the erotetic theory, Inquire, does just this: at any point, reasoners can ask polar questions. These will typically greatly expand the number of alternatives (that is, the size of the question) under consideration, prompting an exponential blowup of alternatives in the worst-case scenario.

To see how polar questions via Inquire help guarantee classical reasoning, we return to the illusory inference from disjunction in (3). Suppose the reasoner asks, after processing the first premise but before processing the second premise, whether Jane is kneeling by the fire or not. While before she was considering a question with two possible answers, Jane-kneeling-and-looking and Mark-standing-and-peering, the addition of the polar question introduces new alternatives. Some of the new alternatives will be ones where Jane is kneeling, others will not.

Now our reasoner updates with the second premise, the putative answer to the question “Jane is kneeling by the fire.” By Q-Update, as informally described above, she will look for possible answers under consideration that overlap with the putative answer. But now, unlike in the case where she did not Inquire, the question is not fully resolved, for the alternatives that remain after update with the partial answer will not agree on other facts about Jane. In particular, she can no longer conclude that Jane is looking at the TV,

for this alternative is not present in all of the answers still under consideration. The fallacious conclusion is therefore blocked. As it turns out, judicious application of this strategy will guarantee classically valid reasoning, though it comes at a potentially very high cost given the multiplication of alternatives that Inquire induces in the worst case.

It is important to note that our account does not posit duality of processes in its explanation of failures and successes of reasoning. Rather, we posit a fixed set of operations on mental representations, and then observe that certain strategies involving those same operations, though costly, will guarantee logical soundness. There is no qualitative change in the kinds of processes at stake.

3.2 On the relation to mental model theory

Mental model theory has offered an account of illusory inferences from disjunction (Johnson-Laird and Savary, 1999; Walsh and Johnson-Laird, 2004), as well as theories of reasoning with quantifiers that are in principle general enough to look beyond syllogisms (Khemlani and Johnson-Laird, 2012; Khemlani et al., 2015). The erotetic theory of reasoning (Koralus and Mascarenhas, 2013) is continuous with work in the mental models tradition. In particular, we follow mental model theory in taking it that reasoning proceeds by updating an integrated mental representation of alternative possibilities in light of successive premise statements. Sentences are interpreted in the context of questions, represented in the form of mental models (Koralus, 2012).

Moreover, mental model theory also has a non-standard semantics for disjunctions that is at least superficially related to the erotetic theory's interpretive commitments. On mental model theory, an atomic statement (as well as a conjunction of atomic statements) will give rise to a single mental model of what is explicitly stated in the sentence being interpreted. But a disjunctive statement, say of two atomic statements, will instead give rise to a set of mental models, one of the first disjunct and one of the second. This is certainly germane to our own approach, which can, in this particular respect, be seen as a reimagining of mental model theory's interpretation of disjunctions. While mental model theory simply postulated an interpretation for disjunctions that involved alternatives, we provide a rationale for that semantics for disjunction, motivated by independent proposals from linguistics and philosophy, namely the notion that disjunctions present questions and therefore alternative propositions. The fact that the erotetic theory incorporates this insight as a central tenet in its approach to mental models is what allows it to recognize the important role that question asking and answering has in human reasoning.

Due to this superficial similarity in the semantics of disjunction, other, weaker similarities arise between the two theories. In particular, mental model theory posits a "matching" operation, responsible for the illusory inference from disjunction in Walsh and Johnson-Laird's (2004) account. On the erotetic theory, Q-Update is in charge of an analogous duty, but, as before, the new lens of question asking and answering offers a new perspective on what the underlying process is doing, what it is a special case of, and why it exists in the first place.

Consequently, mental model theory can incorporate our data from experiment 1 (with indefinite quantifiers) without much difficulty, given the similarities between our Q-Update and the matching procedure. There still exist clear differences in empirical

predictions however, for the effect of premise order we found in experiment 1 is unexpected on mainstream mental model theory's view of matching.

3.3 Contrasting other approaches

The erotetic theory of reasoning predicts and explains the novel patterns of fallacious inferences with quantifiers while enriching a mental model theory account with better-motivated mechanisms for interpreting disjunctive/indefinite sentences and for combining premises. But this is not to say that other theories of reasoning are in principle incapable of incorporating these data. As a reviewer points out, in the probability heuristics model the meaning of "Some A are B " is that $0 < P(B | A) < 1$. Now, if reasoners take $P(B | A)$ to be higher than the threshold for assenting to conclusions in a yes-no task, the probability heuristics model will predict the attractiveness of this inference. The key issue for this account will be to show that indeed those reasoners that draw the fallacious inference consider that the probability of B conditional on A is higher than an independently assessed threshold. We consider this a live possibility that merits more study, but we believe there is reason to think a simple probabilistic account along these lines will face important challenges.

To illustrate, consider again the simpler case of illusory inferences from disjunction as in (13), a schematization of a simplified version of the original inferences from Walsh and Johnson-Laird (2004).

- (13) P₁: $(a \text{ and } b) \text{ or } c$
 P₂: a
 Concl.: Does b follow?

Now, in a fashion similar to the indefinites case, a statement of a in the second premise will raise the probability of b with respect to the first premise; that is $P(b | (a \text{ and } b) \text{ or } c) < P(b | (a \text{ and } b) \text{ or } c, a)$. However, the probability of c is also almost certainly rather high. In fact, assuming that the atomic propositions are independent and each has probability 1/2, we have that

$$P(b | (a \text{ and } b) \text{ or } c, a) = \frac{2}{3} = P(c | (a \text{ and } b) \text{ or } c, a) .$$

Assume that 2/3 is above the relevant threshold, a move necessary to explain the acceptance of the fallacious conclusion b in this view. Then the probability heuristics model would lead us to think that reasoners should be as persuaded that c follows as they are that b follows. That would be very surprising indeed.

Another prima facie promising line of inquiry is to propose a pragmatic account of our data. For the case of illusory inferences from disjunction, the first author of this article has developed one such account (Mascarenhas, 2013, 2014) and shown how a modern theory of scalar implicature, crucially going beyond the means of Grice's original theory, predicts a pragmatically strengthened interpretation of the first premise that ensures the conclusion follows validly. This account can be extended to illusory inferences with universal quantifiers as in experiment 2, but not to the pattern with indefinites established in experiment 1. Informally, the pragmatic strengthening required to make our indefinite inferences classically valid would make premise 1 of (2) mean "There is exactly one pilot, and she writes poems." Not only is this reading not predicted to be a pragmatic

strengthening by any theory of implicature we are acquainted with, it is in fact an extremely implausible interpretation of the indefinite quantifier. While specific readings of “some” are common, amounting to “There is a pilot who writes poems and I have a particular such pilot in mind,” by no means does it follow from such specific readings that there is only one pilot, or that there is only one pilot-poet. On the contrary, use of the indefinite strongly implies, on anyone’s theory of pragmatics, that there are at least two pilots. Consequently, a pragmatic account of our inferences with indefinites is not readily available.

3.4 Concluding remarks

We showed that there are robust fallacious inferences with quantifiers that do not fall under the rubric of syllogisms. These kinds of inferences had previously been overlooked by the reasoning literature. We sketched how the novel inferences were predicted by the erotetic theory of Koralus and Mascarenhas (2013) along with known illusory inferences from disjunction. This insight could only be articulated, and these data investigated, thanks to the consideration of independent accounts of the meanings of disjunctive and indefinite expressions from linguistic semantics. Our results demonstrate the importance of exploring convergences between linguistic theories of meaning and psychological theories of how (representations of) meanings are manipulated to derive novel conclusions (Mascarenhas, 2014).

References

- Bucciarelli, M. and Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, 23:247–303.
- Chater, N. and Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38:191–258.
- Ciardelli, I. A. (2009). Inquisitive semantics and intermediate logics. Master’s thesis, University of Amsterdam.
- Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition*, 54:1–71.
- Frege, G. (1879). *Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle.
- Geurts, B. (2003). Reasoning with quantifiers. *Cognition*, 86:223–251.
- Groenendijk, J. (2008). Inquisitive Semantics: Two possibilities for disjunction. ILLC Prepublications PP-2008-26, ILLC.
- Guyote, M. and Sternberg, R. (1981). A transitive-chain theory of syllogistic reasoning. *Cognitive Psychology*, 13:461–525.
- Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.

- Johnson-Laird, P. N. and Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition*, 71(3):191–229.
- Khemlani, S. and Johnson-Laird, P. N. (2012). Theories of the syllogism: a meta-analysis. *Psychological Bulletin*, 138:427–457.
- Khemlani, S., Lotstein, M., Trafton, J., and Johnson-Laird, P. (2015). Immediate inferences from quantified assertions. *The Quarterly Journal of Experimental Psychology*, 68(10):2073–2096.
- Koralus, P. (2012). The open instruction theory of attitude reports and the pragmatics of answers. *Philosopher's Imprint*, 12(14).
- Koralus, P. and Mascarenhas, S. (2013). The erotetic theory of reasoning: bridges between formal semantics and the psychology of deductive inference. *Philosophical Perspectives*, 27:312–365.
- Kratzer, A. and Shimoyama, J. (2002). Indeterminate pronouns: the view from Japanese. In *Third Tokyo Conference on Psycholinguistics*.
- Mascarenhas, S. (2009). Inquisitive semantics and logic. Master's thesis, ILLC.
- Mascarenhas, S. (2011). Licensing by modification: the case of positive polarity pronouns. In Guevara, A. A., Chernilovskaya, A., and Nouwen, R., editors, *Proceedings of Sinn und Bedeutung 16*, pages 417–429.
- Mascarenhas, S. (2013). An interpretation-based account of illusory inferences from disjunction. Ms. NYU.
- Mascarenhas, S. (2014). *Formal Semantics and the Psychology of Reasoning: Building new bridges and investigating interactions*. PhD thesis, New York University.
- Parrott, M. and Koralus, P. (2015). The erotetic theory of delusional thinking. *Cognitive Neuropsychiatry* (forthcoming).
- Ross, W. (2000). *Aristotle's Prior and Posterior Analytics: A Revised Text with Introduction and Commentary*. Oxford University Press.
- Walsh, C. and Johnson-Laird, P. N. (2004). Coreference and reasoning. *Memory and Cognition*, 32:96–106.
- Wetherick, N. and Gilhooly, K. (1995). "Atmosphere," matching, and logic in syllogistic reasoning. *Current Psychology*, 14:169–178.
- Whitehead, A. N. and Russell, B. (1910). *Principia Mathematica*. Cambridge: Cambridge University Press.