

Illusory inferences in reasoning with quantifiers

(short title at 36 characters: Illusory inferences with quantifiers)

Word count: 2,857

Salvador Mascarenhas and Philipp Koralus
Laboratory for the Philosophy and Psychology of Rationality and Decision
Faculty of Philosophy, University of Oxford

September 1, 2015

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 illusory inferences with quantifiers that do not constitute syllogisms. The inference in (1) is an example of these illusory inferences. Inferences of this sort were accepted by over 39% of subjects in two experiments we report on in this article, very significantly more ($p < .01$) than other invalid controls.

- (1) P_1 : Some pilot writes poems.
 P_2 : John is a pilot.
 Concl.: John writes poems.

Even more strikingly, the inference in (2), with universal quantifiers, was accepted by over 80% of subjects.

- (2) P_1 : Mary has met every king or every queen of Europe.
 P_2 : Mary has met the king of Spain.
 Concl.: *Does it follow that Mary has met the king of Belgium?*

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 (1) 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) P_1 : 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.
 P_2 : 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 is fallacious 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 mental models theory, including the variant designed to account for quantified reasoning (see for example Khemlani and Johnson-Laird, 2012), makes no predictions for the novel inference exemplified in (1). This is largely because (1) is not a syllogism in the traditional sense, and thus lies outside the empirical domain carved out by theories that focus on traditional syllogisms.

By contrast, the erotetic theory of reasoning of Koralus and Mascarenhas (2013) explains the inference in (3) *and* predicts that indefinite quantifiers should give rise to similar fallacies. We begin with a very brief introduction to the erotetic theory of reasoning of Koralus and Mascarenhas (2013) in section 2. As far as we can see, the erotetic theory is the only theory that makes predictions for the quantified cases we consider in this article.

In section 3 we present and discuss three experiments designed to check the predictions of the erotetic theory for these novel quantified cases.

2 The erotetic theory of reasoning

2.1 The theory in a nutshell

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

(4) **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 (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). But on our account, by default, this process of updating proceeds by treating successive premises as questions and maximally strong answers to them. This is Part I of the erotetic principle (4).

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* (Mascarenhas, 2009). 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 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 a sense, premise 1 of (3) can be seen as asking a question: “are we in a Jane kneeling and looking 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. Koralus and Mascarenhas (2013) cash out this notion in terms of propositional overlap: 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. Following the erotetic principle, 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.” From here, the observed fallacious conclusion that Jane is looking at the TV follows immediately.

The theory is also concerned with the problem of success for reasoning, which is ad-

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).

This entirely informal description of the basic tenets of the erotetic theory will suffice for the purposes of this article, but it is important to remark that Koralus and Mascarenhas (2013) give a fully spelled-out formalization of the theory. The formal system we define accounts for a host of existing empirical data points from the literature on reasoning with propositional connectives, 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.

2.2 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 (5-a) shares with (5-b) the ability to raise a question.

- (5) 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 (5-a) with a categorical (i.e. non-disjunctive and non-indefinite) premise that partially overlaps with exactly one of the possible answers to (5-b), giving rise to a fallacious conclusion. The following is a good candidate.

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

On this view, premise 1 of (6) 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 parsed 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 (6) does not constitute an Aristotelian syllogism, and thus has not been considered by extant psychological theories of reasoning with quantifiers.

3 Illusory inferences with quantifiers

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

3.1 Pilot experiment — familiar terms

In a pilot experiment, we tested the two patterns in (7):

- (7) 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 (7-a).

3.2 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 illusory 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.

3.2.1 Design and participants

We recruited 977 workers from the mTurk crowdsourcing platform to solve twelve reasoning problems. 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. 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.

3.2.2 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 (8) we give two representative content variations we used for canonical targets, and in (9) two examples of invalid controls.

- (8) 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?
- (9) 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?

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.

3.2.3 Results

We found that 40% of subjects committed the predicted fallacy. This was a significantly higher rate of reasoning error than in our controls (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

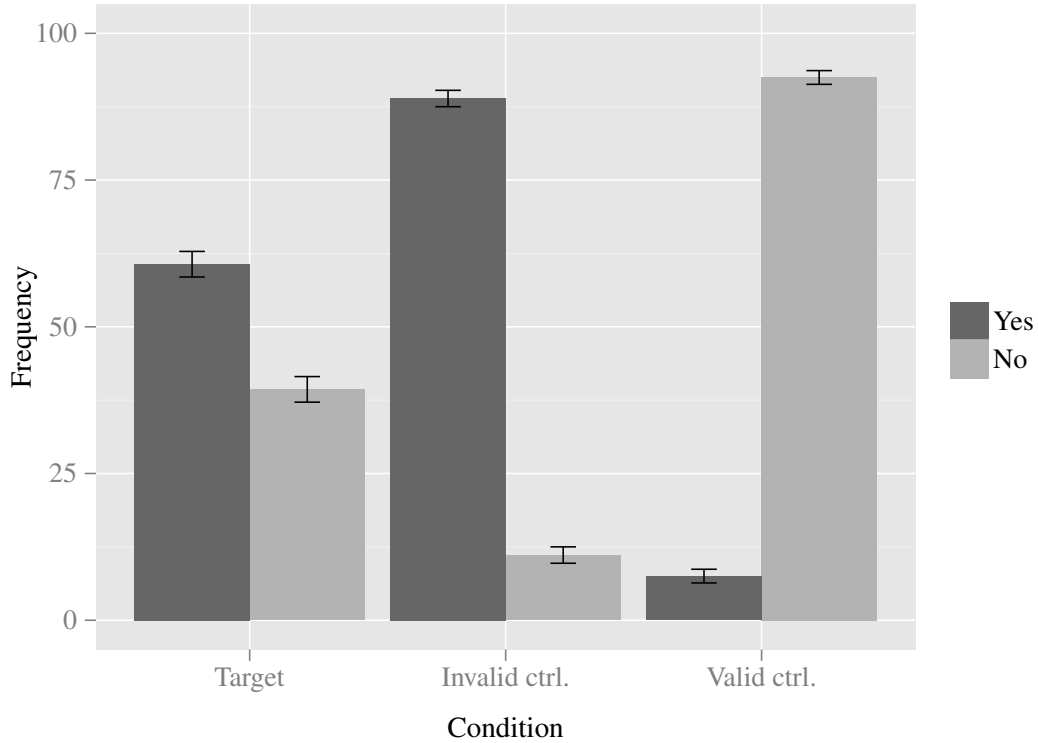


Figure 1: Experiment 1 main results (aggregated targets, invalid controls, and valid controls), with 95% confidence intervals

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 ($\chi^2 > 15.5$, $p < .02$).

We also found a significant order effect on the premises: when the order of the premises in stimuli as in (9) was inverted, we found significantly fewer fallacies for target stimuli. The drop in acceptance was in the order of 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 ($.17 < p < .41$).

3.2.4 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 crucially do not constitute syllogisms. They lie outside the scope of extant theories of reasoning with quantifiers that we are acquainted with (see for example the comprehensive literature review by Khemlani and Johnson-Laird, 2012), but 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 (9) 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. This is not to say that it should be impossible to get a fallacy. 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, mental model theory, even if somehow generalized beyond the special case of traditional syllogisms, still lacks a procedure for combining premises that is relevantly asymmetric.

3.3 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 (10-a) and (10-b) below are representative examples. As with the patterns with indefinites surveyed above, 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 ($N = 203$). 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 $p < .0001$). This result demonstrates that the range of illusory inferences with quantifiers is not limited to indefinite expressions.

4 Conclusions

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. We believe that 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.
- 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.
- Khemlani, S. and Johnson-Laird, P. N. (2012). Theories of the syllogism: a meta-analysis. *Psychological Bulletin*, 138:427–457.
- 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. (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.

Appendix — materials

A Pilot experiment — indefinites with familiar terms

A.1 Procedure

120 subjects recruited from Amazon’s mTurk crowdsourcing platform (45% female). Subjects saw a page with instructions giving examples and one valid and one invalid inference with quantifiers unrelated to indefinites. Each subject then answered three reasoning problems (target, valid control, and invalid control), stating whether a proposed conclusion followed from the premises or not. Which target/control subjects saw was selected randomly. The order of presentation was also randomized. Subjects filled out a brief demographic questionnaire at the end of the study.

A.2 Materials

A.2.1 Targets

- (12) Some student smokes.
John is a student.
Does it follow that John smokes?
- (13) A certain student smokes.
John is a student.
Does it follow that John smokes?
- (14) Some pilot writes poems.
John is a pilot.
Does it follow that John writes poems?
- (15) A certain pilot writes poems.
John is a pilot.
Does it follow that John smokes?

A.2.2 Valid controls

- (16) Every gardener owns a bicycle.
Bill is a gardener.
Does it follow that Bill owns a bicycle?
- (17) Every professor plays soccer.
Mike is a professor.
Does it follow that Mike plays soccer.

A.2.3 Invalid controls

- (18) Few policemen climb mountains.
Frank does not climb mountains.
Does it follow that Frank is a policeman?
- (19) Few custodians swim.
Ted swims.
Does it follow that Ted is a custodian?

A.2.4 Demographic questionnaire

NB: subjects were allowed to decline to answer any of these questions, with no reflection on remuneration.

- (20) What is your college-level background in philosophy, linguistics, and logic?
- a. None
 - b. 1 undergraduate-level course
 - c. 2 or more undergraduate-level courses
 - d. 1 or more graduate-level courses
- (21) How old are you?
- a. Less than 18
 - b. 18 to 29
 - c. 30 to 49
 - d. 50 to 69
 - e. 70 or older
- (22) Are you male or female?
- a. Male
 - b. Female
- (23) How much did you rely on notes or diagrams you made while deciding what conclusions follow?
- a. Not at all
 - b. Somewhat
 - c. A lot

B Experiment 1 — indefinites with unfamiliar terms

B.1 Procedure

We recruited 977 workers from the mTurk crowdsourcing platform to solve twelve reasoning problems (52% male). Statements in each reasoning problem were about biological organisms and properties. 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.

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. 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.

Subjects were shown one inference at a time, each inference consisting of two premises and a proposed conclusion. They answered whether the conclusion followed from the premises. Subjects were also asked not to make notes or use search engines while performing the task. We also asked in the general demographic questionnaire, after solving all reasoning problems, whether subjects had used any search engines. They were told that the experiment would last approximately 5 minutes but were given 10 minutes to answer the questionnaire.

B.2 Materials

B.2.1 Targets

- (24) Some firmicute produces endospores.
Clostridium is a firmicute.
Does it follow that Clostridium produces endospores?
- (25) Some thermotogum stains gram-negative.
Maritima is a thermotogum.
Does it follow that Maritima stains gram-negative?
- (26) Aurantiacus is a chloroflexus.
Some chloroflexus is a monoderm.
Does it follow that Aurantiacus is a monoderm?
- (27) Aeolicus is an acquifex.
Some acquifex is an autotroph
Does it follow that Aeolicus is an autotroph.
- (28) Some bacterium uses all saccharides.
Fusobacterium uses disaccharides.
Does it follow that Fusobacterium uses oligosaccharides?
- (29) Some bacillus metabolizes every sulfide in its environment.
Heliobacillus metabolizes hydrogen sulfide.
Does it follow that Heliobacillus metabolizes dimethyl sulfide?

B.2.2 Valid controls

- (30) Every archaeon has a nucleus.
Halobacterium is an archaeon.

Does it follow that Halobacterium has a nucleus?

- (31) Marina is a nitrospirum.
Every nitrospirum has a vibroid morphology.
Does it follow that Marina has a vibroid morphology?
- (32) Pilosicoli is a brachyspirum.
Pilosicoli is not hemolytic.
Does it follow that some brachyspirum is not hemolytic?

B.2.3 Invalid controls

- (33) Some dictyoglomus is thermophobic.
Turgidum is not a dictyoglomus.
Does it follow that Turgidum is thermophobic?
- (34) Pectinis is not a spirochetes.
Some spirochetes undergoes transverse fission.
Does it follow that Pectinis undergoes transverse fission?
- (35) Some diplocalium contains isochores.
Indicum is not a diplocalium.
Does it follow that Indicum contains isochores?

B.2.4 Demographic questionnaire

NB: subjects were allowed to decline to answer any of these questions, with no reflection on remuneration.

- (36) How old are you?
- (37) What is your college-level background in philosophy, linguistics, and logic?
- a. None
 - b. 1 undergraduate-level course
 - c. 2 or more undergraduate-level courses
 - d. 1 or more graduate-level courses
- (38) What is your college-level background in biology?
- a. None
 - b. 1 undergraduate-level course
 - c. 2 or more undergraduate-level courses
 - d. 1 or more graduate-level courses
- (39) Are you male or female?
- a. Male
 - b. Female

- (40) How much did you rely on notes or diagrams you made while deciding what conclusions follow?
- a. Not at all
 - b. Somewhat
 - c. A lot
- (41) How much did you rely on information you got from search engines when deciding what conclusions follow?
- a. Not at all
 - b. Somewhat
 - c. A lot

C Experiment 2 — universals

C.1 Procedure

200 subjects recruited from Amazon’s mTurk crowdsourcing platform (35% female). Each subject answered 12 reasoning problems, stating whether a proposed conclusion followed from the premises or not. The order of presentation was randomized. Subjects filled out a brief demographic questionnaire at the end of the study.

C.2 Materials

C.2.1 Targets

- (42) Every boy or every girl came to the party.
John came to the party.
Does it follow that Bill came to the party?
- (43) 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?
- (44) Joe hates every rock band or every classical orchestra.
Joe hates the Beatles.
Does it follow that John hates the Rolling Stones?
- (45) Every child or every dog was making a lot of noise.
Bill’s son was making a lot of noise.
Does it follow that Ann’s daughter was making a lot of noise?
- (46) Every sports brand or every computer brand saw an increase in profits this quarter.
Adidas saw an increase in profits this quarter.
Does it follow that Nike saw an increase in profits this quarter?
- (47) Kim knows every father or every mother at this school.
Kim knows Jackie’s father.

Does it follow that Kim knows Claire's father?

C.2.2 Valid controls

- (48) Every teacher in this school has tenure.
Sam is a teacher in this school.
Does it follow that Sam has tenure?
- (49) Every singing contestant played the piano.
Dan is a singing contestant.
Does it follow that Dan played the piano?
- (50) Bill is a taxi driver.
Bill does not have children.
Does it follow that some taxi drivers do not have children?

C.2.3 Invalid controls

- (51) Some teachers are about to retire.
Eva is not a teacher.
Does it follow that Eva is about to retire?
- (52) Jane is not German.
Some Germans follow soccer.
Does it follow that Jane follows soccer?
- (53) Some politicians are corrupt.
Henry is not a politician.
Does it follow that Henry is corrupt?

C.2.4 Demographic questionnaire

As in Experiment 1 (section A.2.4).

List of Figures

1	Experiment 1 main results (aggregated targets, invalid controls, and valid controls), with 95% confidence intervals	7
---	---	---