COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 48 (2024) e13485

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13485

Reasoning From Quantified Modal Premises

Ana Cristina Quelhas,[a] Célia Rasga,[a,b] P. N. Johnson-Laird[c,d]

[a]*Applied Psychology Research Center, Capabilities & Inclusion, ISPA-Instituto Universitario*
[b]*Department of Health Promotion, Instituto Nacional de Saúde Doutor Ricardo, Jorge*
[c]*Department of Psychology, Princeton University*
[d]*Department of Psychology, New York University*

Received 7 September 2023; received in revised form 6 June 2024; accepted 26 July 2024

Abstract

Quantified modal inferences interest logicians, linguists, and computer scientists, but no previous psychological study of them appears to be in the literature. Here is an example of one:

All those artists are businessmen. Paulo is possibly one of the artists. What follows?

People tend to conclude: *Paulo is possibly a businessman* (Experiment 1). It seems plausible, and it follows from an intuitive mental model in which Paulo is one of a set of artists who are businessmen. Further deliberation can yield a model of an alternative possibility in which Paulo is not one of the artists, which confirms that the conclusion is only a possibility. The snag is that standard modal logics, which deal with possibilities, cannot yield a particular conclusion to any premises: Infinitely many follow validly (from any premises) but they do not include the present conclusion. Yet, further experiments corroborated a new mental model theory's predictions for various inferences (Experiment 2), for the occurrence of factual conclusions drawn from premises about possibilities (Experiment 3) and for inferences from premises of modal syllogisms (Experiment 4). The theory is therefore plausible, but we explore the feasibility of a cognitive theory based on modifications to modal logic.

Keywords: Dual processes; Mental models; Modal logics; Syllogisms; Quantifiers; Possibilities

> The negation of "possible not to be" is "not posssible not to be." That is why "possible to be" and "possible not to be" may be thought actually to follow from one another.
>
> Aristotle (*De Interpretatione*, 21b34-6)

Correspondence should be sent to Ana Cristina Quelhas, ISPA-Instituto Universitário, Applied Psychology Research Center, Capabilities & Inclusion Rua Jardim do Tabaco, 34,1149-041 Lisbon, Portugal. E-mail: cquelhas@ispa.pt

1. Introduction

An example of a quantified modal assertion is: "Some plausible arguments may be invalid." It contains the quantifier, *some*, in its grammatical subject: "some plausible arguments," and a modal auxiliary verb, *may*, referring to a possibility in its predicate: "may be invalid." Inferences from such assertions are ubiquitous in everyday life, and we invite readers to draw their own conclusions from these premises:

Some plausible arguments may be invalid.
All invalid arguments are dangerous.
What follows?

People tend to draw this sort of conclusion, where "∴" stands for "therefore":

∴ *Some plausible arguments may be dangerous.*

Such inferences raise a major question for cognitive scientists: What are the underlying mental processes that lead individuals to such conclusions? One answer is that they rely on mental models, which have a long history in studies of inferences based on quantifiers but not possibilities (e.g., Johnson-Laird & Steedman, 1978), and a short history in studies of inferences based on possibilities but not quantifiers (e.g., Ragni and Johnson-Laird (2020). The studies of quantifiers culminated in Khemlani and Johnson-Laird's (2022) program for drawing categorical conclusions rather than those about possibilities, for predicting the differences from one inference to another and for simulating differences in accuracy from one individual to another. Indeed, the two most robust results in studies of quantified reasoning are that such differences in ability are almost as large as they could be, for example, from 15% to 85% correct conclusions in a reported study (see Johnson-Laird, 1983, p. 118 et seq.), and that the difference in difficulty from one sort of quantified inference to another is as large as it could be, for example, from 0% to 100% correct conclusions (ibid.).

Mental models of assertions about possibilities began with a theory (Johnson-Laird & Ragni, 2019) and a set of corroboratory experiments (Ragni & Johnson-Laird, 2020). The present article presents a new synthesis in a theory of quantified modal reasoning— henceforth, the "model" theory. Yet, if you asked scientists at large how individuals make such inferences, most of them are likely to say: on the basis of some sort of logic. But, which "modal" logic (see, e.g., Hughes & Cresswell, 1996)? There is in principle a countable infinity of them (see Johnson-Laird & Ragni, 2019). Yet, most of them are psychologically implausible because they reject both of Aristotle's mutual inferences in the epigraph to our paper.

Psychological studies of modals have investigated several aspects of them, notably children's understanding of the concept of possibility (e.g., Inhelder & Piaget, 1958; Piéraut-Le Bonniec, 1980; Shtulman & Phillips, 2018). A few experimental studies have examined inferences from *categorical* premises—those that make factual claims—to conclusions about possibilities (e.g., Evans, Handley, Harper, & Johnson-Laird, 1999; Hinterecker, Knauff, & Johnson-Laird, 2016). The late Daniel Osherson (1976) pioneered the investigation of inferences from modal premises, and he based his theory on a modal logic known as system T. But, as he conceded, his results were not wholly satisfactory (ibid., p. 232), and his experiments did not examine quantifiers. A recent link to modal logics, however, is psychological studies that invoke their "possible worlds" semantics (e.g., Skovgaard-Olsen, Collins, & Klauer, 2023; but cf. Johnson-Laird & Ragni, 2023).

The present article begins with the model theory (Section 2), which leads to the new theory of quantified modal inferences (Section 3). It then reports four experiments that corroborated the theory's main predictions (Sections 4 through 7). It outlines how in principle a cognitive theory might be formulated to embody a modal logic, and it makes suggestions about how to solve problems that the experiments raise (Section 8). It concludes with an assessment of the model theory and its potential alternatives and a synopsis of those aspects of quantified modal reasoning yet to be investigated (Section 9).

2. The gist of the model theory

The model theory has developed over nearly half a century (e.g., Byrne, 2005; Johnson-Laird, 1975; Johnson-Laird & Byrne, 1991; Khemlani & Johnson-Laird, 2022). At first, it was consistent with classical logics. But, in recent years, experimental findings have forced it to diverge from them. Current versions of the theory deal with modal reasoning without quantifiers (Johnson-Laird & Ragni, 2019) and with quantified reasoning without modals (Khemlani & Johnson-Laird, 2022). The present article unifies these two accounts in order to explain quantified modal reasoning.

The fundamental assumption of the model theory is that if reasoners know the meanings of premises then they can envisage the situations to which they refer, making mental models of them, from which they can draw conclusions. They can do so without the need for a logic or its formal rules of inference. Crucial experiments showed that such rules almost certainly do not underlie human inferences (see, e.g., García-Madruga, Moreno, Carriedo, Gutiérrez, & Johnson-Laird, 2001; Johnson-Laird & Byrne, 1991). Reasoners aim to draw conclusions that are parsimonious, that assert something that is new, and that follow necessarily from the premises, and, failing that, sometimes those that follow as possibilities. Any such theory of what the mind computes is speculative even if it has accrued supporting evidence: it needs to be computable and not to impose an intractable load on working memory. Hence, several programs have implemented the model theory. One of them, mReasoner, implements quantified reasoning but not with possibilities, and it copes both with quantifiers that range over individuals, such as *some* and *all*, and with higher-order quantifiers that in effect range over sets of individuals, such as *few* and *most* (Khemlani & Johnson-Laird, 2022). The program's source code is accessible from https://github.com/skhemlani/mReasoner. Its underlying principle is that reasoners can represent a set of entities by envisaging a small number of exemplars. Another program due to Ragni and Guerth implements modal reasoning with connectives, such as *and*, *or*, and *if*, but not with quantifiers. Its source code is at: https://github.com/CognitiveComputationLab/cogmods/tree/master/modal/2019_guerth/models.

Five assumptions about models, possibilities, and validity, are critical for both the preceding programs and the new model theory integrating quantifiers and modality, and so we spell out each of them in what follows.

2.1. Every model has an iconic structure insofar as possible

Mental models based on visual perception are iconic, that is, their structure represents the structure of the scene, and models based on descriptions are iconic too insofar as possible (e.g., Johnson-Laird, 1983, p. 419, et seq.) unlike Craik's (1943) account. Certain concepts, however, such as negation are abstract. They have no iconic representations, and so they are represented as a symbol linked to its semantics. The models featured in this article are static, but many others are kinematic (Johnson-Laird, 1983, p. 423), such as those for causal relations and sequences of events (Khemlani, Mackiewicz, Bucciarelli, & Johnson-Laird, 2013). Models represent the situations to which descriptions refer, but according to a fundamental principle they depend on representations of the meanings of the descriptions. These meanings and relevant knowledge govern the initial construction of a model and any subsequent modification, its verification, and its use in making inferences (Johnson-Laird, 1983, p. 249).

2.2. Models represent possibilities in default of knowledge to the contrary

Each model represents what is common to different realizations of the same possibility (Johnson-Laird & Ragni, 2019). For instance, the possible outcomes of a coin toss have three models (head, tails, neither). The category "neither" allows for various mishaps, such as that the coin lands vertically in a furrow. Likewise, according to the model theory, the assertion

The coin toss came up heads or tails

refers to a conjunction of two possibilities:

Possibly the coin toss came up heads and possibly the coin toss came up tails.

Each possibility holds in default of knowledge to the contrary. And so they can be withdrawn without contradiction, but if both are refuted, then the disjunction is false. A primordial "exclusory" inference uses the exclusion of one possibility to infer the other. Even 3-year-old children can make this inference, and perhaps non-human primates can do so too (see, e.g., Feiman, Mody, & Carey, 2022, for a review, and for their own ingenious studies with infants). Another sort of inference from disjunctions corroborates the semantics of a conjunction of possibilities. People infer such possibilities from disjunctions (Hinterecker et al., 2016). So, they accept an inference from the preceding disjunction such as

∴ *It is possible that the coin came up heads.*

This inference is not valid in classical logic, where a *valid* inference is definable as one whose conclusion is true in every case in which all its premises are true (Jeffrey, 1981, p. 1). A valid inference therefore has no counterexample in which its premises are true, but its conclusion is false. The preceding inference has a counterexample in which it is impossible that the coin came up heads but possible that it came up tails, so the disjunctive premise is true, but the modal conclusion is false.

Just as a model of a possibility can be withdrawn without contradiction, so too reasoning is defeasible, that is, "nonmonotonic" in artificial intelligence (AI), in that individuals can withdraw any conclusion (e.g., Marek & Truszynski, 2013). Unlike AI systems, however, the model theory postulates that individuals search for knowledge to resolve an inconsistency between a conclusion and evidence to its contrary (Khemlani et al., 2013)—a process that is also implemented computationally (see Johnson-Laird, Girotto, & Legrenzi, 2004). Knowledge can modulate the interpretation of an assertion, preventing the construction of a model otherwise consistent with the assertion's meaning. Thus, the assertion, *all spouses are married* can trigger the knowledge that blocks a model of the possibility of a married person who is not a spouse. This process reinterprets the assertion as equivalent to *all and only spouses are married* (e.g., Quelhas & Johnson-Laird, 2017).

2.3. Varieties of possibilities

A previous analysis of possibilities in daily usage distinguished three main sorts (Johnson-Laird & Ragni, 2019). *Alethic* modals concern relations between concepts and between propositions, such as the premises and conclusions of inferences. These relations include necessity, possibility, and impossibility. Hence, an inference is alethically *necessary* in case its conclusion holds for one or more possibilities or facts that the premises describe and does not deny any of the others. An assertion such as *It is necessary that it rains* can assert an alethic relation from rain to some other eventuality, such as *for the plants to flourish*. Likewise, *necessarily, it will rain* describes an alethic relation in which some other situation implies that it will rain. Philosophers distinguish between necessary conditions—as in the first of these cases and necessary truths—as in the second of these cases. But, if one asks what it is that makes a truth necessary, the answer in daily life is that it follows from other information, which may be the meaning of the assertion itself, for example, *A person is either married or not married* is necessarily true. The *epistemic* modality concerns the occurrence or non-occurrence of events inferred from empirical knowledge of the world. The modality therefore ranges from impossible, through barely possible, up to certain (Lassiter, 2017; White, 1975). It is non-numerical, but it can underlie numerical probabilities (Johnson-Laird & Ragni, 2019). The *deontic* modality depends on moral principles, rules, regulations, and social conventions, and so, for example, it expresses permissions and obligations, for example, *You may leave now.* Indeed, certain speech acts, if felicitous, can create deontic situations, which are inexpressible as probabilities. Our previous account analyzed cues in syntax and context to these different interpretations and illustrated common ambiguities.

Like Aristotle (see the epigraph), the model theory postulates that an assertion of a possibility presupposes the possibility of its negation. Hence, an assertion about an event

It may happen

presupposes

It may not happen.

If this presupposition is false, then it follows that the event is certain to happen. Naive individuals—that is, those who are innocent of logic—accept inferences of this sort and their converses (Ragni & Johnson-Laird, 2020). They are invalid in those modal logics, such as system T and those systems that include it, because they treat the occurrence of an event as implying that it may happen, which cannot then allow that it may not happen.

Human working memory has a limited capacity, so models are parsimonious, and individuals tend to condense separate possibilities into one, provided that they are consistent with one another (Ragni & Johnson-Laird, 2020). This condensation explains several otherwise perplexing phenomena, which we elide, because they are irrelevant to the present study (see Rasga, Quelhas, & Johnson-Laird, 2022). But they include the paradoxes of "free choice" permission (e.g., Kamp, 1973), which generalize to inferences of *or*-deletion, such as

Eric does drink red or white wine for lunch.
∴ *Eric does drink red wine for lunch.*

As the model theory predicts, participants tend to accept such inferences (Johnson-Laird, Quelhas, & Rasga, 2021; Rasga et al., 2022) but to reject inferences of this sort:

Eric drank red or white wine for lunch today.
∴ *He drank red wine for lunch today.*

The importance of consistency for *or*-deletions is illustrated in inferences based on quantified assertions that do not concern permissions, such as

Some of the customers ate steak or lobster.
∴ *Some of the customers ate steak.*

If *some* is replaced with *all* in the preceding inference:

All of the customers ate steak or lobster.
∴ *All of the customers ate steak.*

then reasoners reject the inference—the proportions of customers who ate steak and who ate lobster are inconsistent with one another. The same contrast occurs between *none* and *one* and *few* and *most* (Rasga et al., 2022).

2.4. Reasoning depends on dual interacting systems

Wason (1960) introduced two systems for reasoning in a "dual process" account, and the first algorithm implementing them was for his well-known selection task, which concerns the selection of evidence to test whether a conditional hypothesis is true or false (Johnson-Laird & Wason, 1970). Many others have taken up Wason's idea (e.g., de Neys, 2022; Evans, 2008), and it was made famous in Kahneman's (2011) distinction between thinking fast and thinking slow. From its beginnings, the model theory recognized the distinction and contrasted implicit inferences, which are rapid and intuitive, with explicit inferences, which are slower and deliberative (Johnson-Laird, 1983, Ch. 6). The intuitive system cannot use working memory to store the results of intermediate computations, so it yields only a single model at a time, and it represents only what is true according to the premises (Khemlani & Johnson-Laird, 2022; Khemlani, Byrne, & Johnson-Laird, 2018). So, the intuitive system makes predictable and compelling errors. For example, given the following sort of inference:

Only one of the following assertions is true:
At least some of the brown beads are not round, or
None of the brown beads are round.
Is it possible that none of the round beads are brown?

most participants responded "yes" (Yang & Johnson-Laird, 2000). According to the theory, they build an intuitive model of the second premise, which supports the conclusion, and so they accept the inference. But deliberation leads them to think about what is false. And, if the second premise is true, the first premise is false, and so *all the brown beads are round* contradicts the second premise and rules out the putative conclusion. Many other experiments have corroborated these "illusory" fallacies drawn from intuitive models. Yet participants do well with the control problems (for a review, see Khemlani & Johnson-Laird, 2017).

The mReasoner program has a loop in deliberation between searching for a model that is a counterexample to an intuitive conclusion and weakening the conclusion so that it holds in all the models of the premises so far constructed (Khemlani & Johnson-Laird, 2022). When it fails to find such a conclusion, it responds: *Nothing follows*—an evaluation that also occurs when it is unable to construct a single model that integrates the premises.

2.5. *The model theory replaces the logical concept of validity with alethic necessity*

In classical logic, a *valid* inference is one whose conclusion is true in every case in which all its premises are true. In contrast, as we pointed out earlier, the model theory's criterion of a correct inference is alethic necessity. An inference is *necessary* if its conclusion describes only one or more possibilities or facts that its premises describe, and does not deny any of them, as in:

It is freezing or snowing, or both.
∴ *It may be freezing.*

If a conclusion describes a possibility outside those that the premises describe, for example:

It is freezing,
∴ *It is freezing or my teeth are chattering, or both,*

its inference is not necessary, and so individuals tend to reject it even though it is valid in standard logics. When the premises are inconsistent with a conclusion, then the conclusion is alethically impossible given the premises. Many inferences are both necessary in the present sense and valid, but the first of the two inferences above is necessary but not valid, and the second of them is valid but not necessary. Nothing in classical logic can justify the withdrawal of a valid inference, not even a direct denial of its conclusion. Which merely forms an inconsistent set of premises when it is added to the premises, and inconsistencies in classical logic yield any conclusion whatsoever including the one that is denied. Their inferences are valid because there can be no case in which the premises are true, and a counterexample calls for true premises and a false conclusion.

In contrast to classical logic, inconsistent assertions in the model theory have only a null model—an empty one that denotes that there are no possibilities that the assertions describe, and the conjunction of the null model with any other model yields only the null model once more. So, inconsistencies have only local consequences: There is something wrong with their premises. They do not yield any other conclusions. Indeed, people do not treat inconsistencies as licenses to infer any conclusions whatsoever (see Johnson-Laird et al., 2004). They are glitches and an intimation that human reasoning diverges from classical logic. We can now consider the new theory of quantified modal assertions.

3. The model theory of quantified modal inferences

The fundamental assumption of the model theory is that reasoners who understand the meanings of a set of premises can construct models of the situations to which they refer and draw conclusions from these models. So, this section describes the theory's semantics for quantified modal assertions (Section 3.1), and how they can be used to construct models from which conclusions follow (Section 3.2). It presents only what readers need to know in order to understand the theory's three main predictions (Section 3.3) that inspired our experimental studies. Those readers who are interested in more details can consult a description of the mReasoner program (as described in Khemlani & Johnson-Laird, 2022).

3.1. The meanings and models of monadic assertions

The model theory postulates that a quantified assertion about properties establishes a relation between the set of entities to which the subject of the sentence refers and the set of entities to which the object of the sentence refers. This idea goes back at least to Boole (1847) and differs from Montague's (1974) elegant treatment of all noun phrases as having uniform semantics, albeit one that is implausible as psychology (see Johnson-Laird, 1983, Ch. 8). The Boolean semantics copes with all quantifiers, including those that call for a high-order logic, such as *more than half*, which quantify not only over individuals but also over sets of individuals. The quantified assertions under investigation are *monadic*, that is, a single quantifier governs a single argument. The mental parser uses the grammatical constituents that it recovers from each premise to compose a representation of its meaning. These meanings can then be used to construct intuitive models of premises, to search for deliberative models, and to constrain the formulation of conclusions. A quantified assertion yields a conjunction of models in which each model represents a set of possible entities that hold in default of knowledge to the contrary. These processes and others are embodied in the mReasoner program. For the premises in the present experiments, a quantified relation is one between sets, but there are many other relations: spatial, temporal, causal, and as in a plethora of verbs such as: *see*, *meet*, *tell*, *marry*, and so forth, which concern particular relations.

In the classical logic for quantifiers, the symbols corresponding to "some artists" imply that artists exist, whereas the symbols corresponding to "all artists" do not imply that artists exist. One consequence is that a syllogism, such as

All artists are Bohemians,
No Bohemians are clinicians,
∴ *Some clinicians are not artists*,

is invalid because both premises can be vacuously true in case the various individuals do not exist, whereas the conclusion makes the claim that certain individuals exist. But, if artists and clinicians do exist in the situation under description, the inference is valid. Two reviewers queried our use in experiments of *those* in noun phrases such as "All those artists." But the reason for this usage goes back to a distinguished logician, Boolos (1984), who pointed out that the preceding inference is not valid, contrary to claims in a psychological paper. The butt of his criticism, Johnson-Laird and Bara (1984) replied: the first premise in their inference was in fact: *All* the *artists are bohemians*, and they had instructed their participants that each inference concerned three sets of individuals gathered together in a room. They also noted that Boolos was neither the first (cf. Kyburg, 1983, p. 266) nor probably the last to err in this way. To solve this problem of vacuous truth, one reviewer suggested that the present experiments could have used noun phrases such as *All the administrators are bankers*, as translated into Portuguese, the language in which we carried out the experiments:

Todos os administradores são banqueiros.

But this sentence is also the translation of *All administrators are bankers*, and so it leaves open whether a particular set of administrators exists. To ensure that there was no doubt about their existence, our experiments therefore used unambiguous noun phrases such as *Todos aqueles administradores* (*All those administrators*). The other reviewer argued that this usage introduced a confound, that it was unnecessary, and that we should have used simple quantified phrases, as in:

All artists are possibly cyclists.

As the reviewer wrote: "Given their participants' background knowledge, these participants would have known that this sentence is not vacuously true: of course there are artists!" But sentences such as the preceding example present a real danger that participants will not believe them—some artists are not cyclists, and beliefs are a notorious source of biased spontaneous conclusions from quantified premises (see, e.g., Bucciarelli & Johnson-Laird, 1999, 2024; Markovits, 2023; Oakhill & Johnson-Laird, 1985). Hence, our preference was to use premises that would avoid both vacuity and the biasing effects of beliefs. We return later (in the General Discussion) to whether noun phrases containing *those* introduce a confound.

The clear-cut distinction in logic between the existential implication of the quantifier corresponding to "some" and its lack for the quantifier corresponding to "all" does not mirror everyday discourse. For instance, the following assertion

Some of the characters in Shakespeare's Henry VI existed and some did not exist

treats existence as a predicate in a way that is outside classical logic (see also Routley, 1982). There are "free" logics that allow existence to be treated as a predicate (see Priest, 2008, Ch. 13), which is sensible for the analysis of everyday reasoning.

In the model theory, the representation of the meaning of a monadic assertion, such as

All those artists are bakers,

Table 1
The model theory's meanings of the six sorts of assertion (stated informally), two concerning an individual's membership of a set and four concerning monadic quantified assertions concerning intersections of sets; the numbers of entities to which the subject refers; and constraints on the numbers of entities in the sets, where | B | denotes the cardinal number of entities in set B (see Boole, 1847, p. 20 et seq.; Khemlani & Johnson-Laird, 2022)

Assertion	Informal Meaning in Terms of Sets	Number of Entities in Subject	Number of Entities in Relevant Sets
a is one of those B	a is a member of B	1	$\lvert B \rvert > 1$
a is not one of those B	a is not a member of B	1	$\lvert B \rvert > 1$
All those A are B	Intersection of A and not B is empty: A is included in B	>1	$\lvert A \& B \rvert = \lvert A \rvert$
Some of those A are B	Intersection of A and B is not empty	>1	$0 < \lvert A \& B \rvert \leq \lvert A \rvert$
None of those A is B	Intersection of A and B is empty	≥ 1	$\lvert A \& B \rvert = 0$
Some of those A are not B	Intersection of A and not B is not empty	≥ 1	$0 < \lvert A \& \neg B \rvert \leq \lvert A \rvert$

Note. Lowercase letters denote particular individuals, capital letters denote sets of individuals (or properties), and "¬" denotes negation.

follows the Boolean analysis, and so it means that the intersection of the set of those artists and the set of non-bakers is empty (see Boole, 1847, p. 21). The model theory's semantics for the preceding sentence also include information relevant to the construction of mental models:

1. The subject of the sentence specifies a set of artists of two or more definite individuals, who exist in the situation under description.
2. The object of the sentence specifies an indefinite set of bakers of at least the same number as those to which the subject refers.
3. The relation is factual and affirmative in that the subject's referents are included within those to which the object refers, so the cardinal number of those who are members of both sets is equal to or less than the number who are in the object's set. Similar principles underlie the meanings of assertions about the membership of sets, such as *Anna is an artist*, and about assertions based on other quantifiers. One set can be a member of another, as in: *Scots are divided about Scottish independence*, and one set can be included in another, as in: *Scots are British*. Given that Phil is a Scot, it follows that he is British, but it does not follow that he is divided about Scottish independence. We pursue the distinction between set-membership and set-inclusion no further here. Table 1 states the model theory of the meanings of a representative set of assertions, the number of entities to which the subject of the assertion refers, and constraints on the number of individuals in each set. We emphasize that the latter are intended, not as a gloss on the meanings, but as guidelines for the construction of mental models.

The denial of assertions is a well-known source of difficulty in everyday reasoning because individuals do not have immediate access to the possibilities to which negative assertions refer. Psychologists who favor "embodied" cognition (e.g., Barsalou, 1999; Glenberg, 2010) have often sought to replace a negative with an equivalent affirmative. This transformation

Table 2
The categorical assertions under investigation and examples of their typical intuitive and deliberative models. Each assertion has other intuitive and deliberative models that vary in their number of tokens and typicality

Assertion	Intuitive Model	Deliberative Model
a is B	a B	No change
	B	
	B	
a is not B	a	a ¬ B
	B	B
	B	B
All those A are B	A B	A B
	A B	A B
	A B	¬A B
Some of those A are B	A B	A B
	A B	A ¬ B
	A	¬A B
None of those A is B	A ¬ B	A ¬ B
	A ¬ B	A ¬ B
	B	¬A B
Some of those A are not B. (Not all those A are B)	A ¬ B	A ¬ B
	A ¬ B	A ¬ B
	A	A B

Note. Lowercase letters denote particular individuals, capital letters denote sets of individuals (or properties), and "¬" denotes negation.

works well for binary predicates, for example, *it is not open* is equivalent to *it is closed*. But it fails in other cases, for example, *Anna is not an artist* has no obvious affirmative complement that covers all the possibilities (Vance & Oaksford, 2021). A corollary is that reasoners need to construct the complement of a set of models. For instance, the negation of a conjunction, such as *It is hot and it is raining*, calls for a conjunction of three possibilities that hold in default of knowledge to the contrary: It is not hot and not raining, it is hot but not raining, and it is not hot but raining. Individuals often overlook a possibility (Khemlani, Orenes, & Johnson-Laird, 2014).

Table 2 presents typical intuitive and deliberative models for two assertions about an individual's membership and non-membership of a set and four monadic quantified assertions about relations between sets. The negation of *All those A are B* is equivalent to *Not all those A are B* or equivalently *Some of those A are not B*, and the negation of *Some of those A are B* is equivalent to *None of those A is B*. Anything that is common to a set of possible individuals follows as a necessary inference about them, for example, *At least some B are A* follows necessarily from the models of *All those A are B*.

3.2. Reasoning about possibilities

The model theory postulates that an assertion about a possibility presupposes the possibility that it does not occur (see Section 2.3). So, the assertion

Possibly all those artists are bakers

presupposes

Possibly not all those artists are bakers.

Because each model represents a possibility, the presupposition has a crucial consequence for the model theory of reasoning about possibilities. It does not require a special process: just build an intuitive model of the occurrence of the possible event and, if elicited, a deliberative model of its non-occurrence. If the two processes yield distinct models, then a conclusion expressing the possibility of the event in the intuitive model is a necessary inference. Here is a simple illustration of the two processes for a pair of syllogistic premises:

Some of those artists may be bakers.
All those bakers are chefs.

First, build an intuitive model of the possibility to which the premises refer but representing only a small number of individuals, for example:

artist	baker	chef
artist	baker	chef
artist		

Reasoners who go no further draw a categorical conclusion:

∴ *Some of those artists are chefs.*

It is wrong, though it would have been a necessary conclusion had both the original premises been categorical. Reasoners who deliberate can build a model of the presupposition of the first premise, that is, its negation: *None of the artists is a baker,* which with the second premise, *All those bakers are chefs,* yield a deliberative model that refutes the categorical conclusion above because it represents none of the artists as chefs:

artist	¬ baker	
artist	¬ baker	
	baker	chef
	baker	chef

Hence, a conclusion of the intuitive possibility follows necessarily:

∴ *Some of those artists may be chefs.*

Of course, there are other deliberative models and other ways to reach the same conclusion, but they need not detain us: Naive reasoners can draw a necessary conclusion about a possibility from models of its occurrence and of its non-occurrence. Intuition can build only a single model, and so the other model calls for deliberation.

The preceding example illustrates the model theory. Those reasoners who consider only the intuitive model will draw an erroneous categorical conclusion. But those who also consider the deliberative model of the possibility's presupposition will draw a necessary conclusion about a possibility. Of course, those reasoners who are unable to build a single integrated model of the premises will infer that nothing follows from the premises.

A necessary categorical conclusion from categorical premises must hold in a deliberative model that is irrefutable. Some of these conclusions also hold in an intuitive model, and so the theory predicts that they will be easier to infer than those about possibilities. Reasoners who do not deliberate can infer the former but not the latter.

The theory also predicts that inferences from affirmative premises are easier—faster and more accurate—than those from negative premises (e.g., Johnson-Laird, 1983; Khemlani et al., 2014; Newstead & Griggs, 1983; Wason, 1959, p. 245). Likewise, the negation of a quantified assertion, such as: *Not all those artists are chefs*, calls for deliberation about what can no longer hold in the models of the original affirmative assertion. There can still be some artists who are chefs provided that at least one of them is not a chef.

Readers can emulate the preceding inferential procedure. Hence, if they wonder about the models of premises that we present, they need only ask themselves whether the intuitive model of premises about a possibility holds for its categorical equivalent and whether its deliberative models hold for the denial of this categorical. If so, then a conclusion expressing the possibility of the event in the intuitive model is a necessary inference. Likewise, for categorical premises, they need only ask themselves whether the conclusion holds in its deliberative model and has no countermodel in which it is false. We have omitted many details from the procedure implemented in mReasoner (see Khemlani & Johnson-Laird, 2022), but they do not affect the following predictions.

3.3. The model theory's three main predictions

The three main predictions are about quantified modal reasoning, and they are as follows:

Prediction 1: Reasoners who construct only an intuitive model of premises about a possibility will draw an erroneous categorical conclusion. Those who also construct the deliberative model of the premises will draw a necessary conclusion about a possibility granted that the two models support one. Those who are unable to integrate the premises into a single model will respond that nothing follows from the premises.

Prediction 2: Reasoners will draw their own necessary conclusions more often from categorical premises that yield them from their intuitive models than from any other sort of premises. These other sorts include those with a premise about a possibility. So, if everything else is the same, categorical premises should yield more necessary inferences than premises about possibilities do.

Prediction 3: Reasoners will infer necessary conclusions from affirmative premises more often than from negative premises because of the difficulty of constructing the complements of sets of models.

Our experiments tested these predictions. They called for the participants to draw their own conclusions from given premises—a procedure that is more informative than any other for a domain that has not been studied before.

4. Experiment 1: Inferences about membership of sets

Our first experiment was an exploratory study to test the model theory's Prediction 1. It used the modal operator *possibly* (translated from the Portuguese, *possivelmente*, the language of all our experiments) in each of four sorts of inferential problem based on pairs of premises from which the participants had to draw their own conclusions. According to the model theory, models of premises vary from one occasion to another, and from one person to another. So, our diagrams are of typical models of the premises rather than atypical ones, but together the intuitive and deliberative models yield necessary conclusions from premises if there are any.

Inference 1 had premises, such as

All those administrators are possibly bankers.
Cynthia is one of those administrators.
What follows?

We used "those" in the first premise to ensure that the participants understood that it referred to administrators who exist in the relevant situation (see Section 3.1 for the reasons). The model theory's Prediction 1 is that when a premise refers to a possibility, its intuitive model represents the possibility as occurring, and its deliberative model represents the presupposition of it not occurring. So, the premises above have an intuitive model in which the situation occurs, that is, all those administrators are bankers:

administrator	banker	Cynthia
administrator	banker	

It yields the erroneous categorical conclusion:

∴ *Cynthia is a banker.*

The premises have a deliberative model in which the situation does not occur, that is, not all the administrators are bankers, where "¬" is the symbol for negation:

administrator	¬ banker	Cynthia
administrator	¬ banker	
administrator		

In the model theory, a single model of a possibility in a set of models suffices for its necessary inference. So, those participants who construct both models should tend to draw the necessary conclusion:

∴ *Cynthia is possibly a banker.*

Inference 2 had premises such as

All those administrators are bankers.
Cynthia is possibly one of those administrators.
What follows?

Its intuitive model represents that Cynthia is one of those administrators, and so it is the same as the one for Inference 1. And its deliberative model represents that Cynthia is not one of those administrators:

administrator	banker	
administrator	banker	
administrator		¬ Cynthia

The two models together yield the same conclusion about a possibility as Inference 1 does.

Inference 3 had premises such as

All those possible administrators are bankers.
Cynthia is an architect.

It yields the same intuitive model as the previous two inferences. But deliberation yields two different construals. One is that if individuals who are possible administrators are bankers, then actual administrators are too. So, this construal yields a deliberative model, which is the same as the intuitive one:

administrator	banker	Cynthia
administrator	banker	

Hence, this construal yields a categorical conclusion that Cynthia is a banker. The other construal is that if individuals who are *possible* administrators are bankers, then those who are actual administrators may not be. This construal yields a deliberative model of the same sort as Inference 1, such as

administrator	¬ banker	Cynthia
administrator	¬ banker	
administrator		

Hence, the intuitive and deliberative model yield the modal conclusion:

∴ *Cynthia is possibly a banker.*

Inference 4 had a pair of premises with no predicate in common, such as

All those administrators are bankers.
Cynthia is possibly a carpenter.

Participants might conclude that Cynthia is possibly a banker, but there are no grounds for such an inference, and so other than drawing a conclusion from an individual premise, which hardly ever happens in studies of reasoning, the model theory predicts that participants should respond that nothing follows from the premises.

4.1. Method

4.1.1. Participants

The experiment tested 102 psychology undergraduates (83 females, 17 males, two non-binaries) from Istituto Superior de Psicologia Applicada in Lisbon. Their mean age was 21.3 years ($SD = 7.9$). Their participation was voluntary, for which they received a course credit. They were native speakers of Portuguese, the language used in all our experiments. The G*power 3.1 program yielded a sample size of 94 in order to detect an effect size (0.35) of correct inferences with 95% power and an alpha error $p = .05$ using a one-tailed Wilcoxon test (Faul, Erdfelder, Buchner, & Lang, 2009).

4.1.2. Design and materials

The participants acted as their own controls, and their task was to draw their own conclusion from each of three instances of the four sorts of inferential problem for a total of 12 trials. The first premise used a universal quantifier to assert that one set was included in another, and the second premise asserted that a named individual was in a set. Table 3 presents the four sorts of inferential problem, their intuitive and deliberative models, and their predicted conclusions.

The 12 contents were from four domains (business, studies, sports, food), and each topic occurred in three sentences. Half of them had female proper nouns as subjects, and half of them had male proper nouns as subjects. There were four versions of the materials in order to rotate the four contents over the four sorts of inference. The Supporting Information at https://osf.io/gpsmk/ presents the sentences (in the four versions) used in this experiment and the remaining experiments in both an English translation and the original Portuguese. The 12 inferences were presented in a different random order to each participant.

4.1.3. Procedure

The participants were each seated at a computer in the same room in groups of about 20. They received a link to access the experiment in a Qualtrics program. They read the instructions, which stated the general nature of the experiment—that it examined reasoning and was not a test of intelligence or personality. They gave their informed consent and then stated their age and gender. In the experiment proper, each inference was presented on a separate screen. Fig. 1 shows a typical trial, which repeats the key instruction about how the participants should draw a conclusion, framed in terms of preserving truth to avoid biasing their performance in favor of the model theory's normative criterion of necessity. The participants were allowed as much time as they needed to complete the experiment, which took them about 25 min.

4.2. Results

Two independent judges classified the participants' conclusions into seven sorts derived from the three in the model theory's Prediction 1 and from those that the participants drew. The judges agreed with one another (Cohen's $k = .98$), and they resolved their few disagreements in discussion. In this experiment and all the subsequent ones, we made

Table 3
Typical intuitive and deliberative models for the four sorts of pairs of premises in Experiment 1 ($N = 102$) and the percentages of the three main sorts of the model theory's predicted conclusions that the participants drew for themselves

The Four Sorts of Premises	Instances of Intuitive and Deliberative Models	Three Conclusions That Participants Tended to Infer
1. *All those A are possibly B.* *c is one of those A*	A B c A ¬B c A B A ¬B A	**∴ c is possibly B: 70** ∴ c is B: 21 Nothing follows: 4
2. *All those A are B.* *c is possibly one of those A*	A B c A B A B A B ¬A c	**∴ c is possibly B: 60** ∴ c is B: 19 Nothing follows: 6
3. *All those possible A are B.* *c is one of those A*	A B c A B c <u>A B A B</u> A B c A B c A B A ¬B c A ¬B	<u>**∴ c is B: 47**</u> **∴ c is possibly B: 28** Nothing follows: 11
4. *All those A are B.* *c is possibly one of those D*	No integrated model of both premises	**Nothing follows: 69** ∴ c is possibly B: 4 ∴ c is B: 2

Note. A and B in models denote sets of individuals such as *doctors* and *surfers*, and c denotes an individual such as *Cynthia*. Conclusions in bold are for necessary conclusions in the model theory, that is, from both intuitive and deliberative models. Inference 3 is open to two different interpretations reflected in its two deliberative models (see text). The balances of percentages for each inference are for miscellaneous conclusions occurring for no more than 10% of its responses.

Please read carefully the two sentences below. You should consider that the sentences are true and from them you should draw a conclusion that must also be true. If you think that it is not possible to conclude anything, please respond: no conclusion.

All those artists are possibly cyclists.

Manuela is one of those artists.

Therefore ____

Fig. 1. A typical trial from the experiment, which repeats the key instruction about drawing a conclusion.

statistical analyses of all the data from the participants, and the raw data are available at https://osf.io/gpsmk/.

Table 3 presents the percentages of conclusions from each of the four sorts of inference, and the balance of percentages was for various miscellaneous conclusions occurring on less than 10% of trials for the relevant inference. The participants drew more of the model theory's predicted conclusions than other conclusions: 68 participants out of 102 did so, 19 participants drew more unpredicted conclusions than predicted ones, and the remaining 15 participants were tied (binomial test, with a prior of 0.5, $p < 1$ in 10 million, corroborating prediction 1). Likewise, as Table 3 shows, the majority of conclusions for each of the four sorts of inference were those that the model theory predicted: 91% for Inference 1; 85% for Inference 2; 86% for Inference 3; and 69% for Inference 4 (Wilcoxon test, $z = 1.83$, $p < .05$).

The model theory predicts that one interpretation of Inference 3 yields a categorical conclusion and thereby reduces the number of conclusions of the sort: *c is possibly B*. We examined the mean number of trials (out of the three for each sort of inference) in which participants drew conclusions about possibilities. Their mean for Inference 3 was 0.83 conclusions out of three, which was reliably smaller than their mean of 1.94 for Inferences 1 and 2 (Wilcoxon test, $z > 6.53$, $p < .001$, effect size: $r = .46$). Very few participants in the experiment drew any probabilistic conclusions, which is consistent with the hypothesis that to elicit them reliably the content or context of an inference needs to refer to probabilities (Johnson-Laird, Khemlani, & Goodwin, 2015, p. 207).

In sum, the experiment corroborated the model theory's Prediction 1: The participants tended to draw a necessary conclusion, where one followed from the premises, or an erroneous categorical conclusion that followed from an intuitive model of the premises, or no conclusion whatsoever for premises that had nothing in common.

5. Experiment 2: A comparison of categorical and modal inferences

The experiment was designed to test the three main predictions of the model theory. It included Aristotle's two fundamental inferences in which the first premise is affirmative, *All those A are B*, and the second premise either asserts that an individual is an *A* or asserts that an individual is not *B* (see *Prior Analytics*, Book 1, 19, in Barnes, 1984). It included analogous inferences in which the first premise was negative, *None of those A is a B*. For each of these four categorical inferences, it also examined modal versions in which the second premise for each of them concerned a possibility, for example, *c may be one of those A*.

5.1. Method

5.1.1. Participants

This experiment tested a new sample of 91 psychology undergraduates (78 females, 13 males) from the same population as before. Their mean age was 20.7 years ($SD = 6.2$). The G*power 3.1 program yielded a sample size of 94 in order to detect an effect size (0.35) of correct inferences with 95% power and an alpha error $p = .05$ using a one-tailed Wilcoxon test.

5.1.2. Design, materials, and procedure

The participants acted as their own controls and their task was to draw their own conclusions from eight sorts of pairs of premises. The eight sorts were a result of manipulating three variables: the subject of the first premise was either a universal affirmative (*All those A*) or negative (*None of those A*), the second premise asserted either that an individual was in set *A* or not in set *B*, and it was either categorical or about a possibility. Table 4 presents the eight sorts of inference, their typical intuitive and deliberative models, and the percentages of the different conclusions that the participants drew.

Each participant had two instances of the eight sorts of premise for a total of 16 inferences. The first premise had contents referring to everyday occupations (e.g., doctors) and avocations (e.g., surfers), and the second premise had a male or female name as its subject and one of the two topics in the first premise in its predicate. The contents were assigned twice at random to the eight sorts of premise. Each participant was assigned one of these two orders at random. The Supporting Information presents the full set of contents in English and in their original Portuguese.

The procedure was similar to the one for Experiment 1, with a Qualtrics program presenting the instructions and the 16 trials, and there was no limit on the time for the participants to respond. The presentation of each pair of premises was identical to the one in Experiment 1 (see Fig. 1). After the 16 inferences, there were two checks on the participants' attention. One was a simple common-sense inference, and the other asked whether the participant had any difficulty in paying attention to the problems.

Table 4
Typical intuitive and deliberative models for the eight pairs of premises in Experiment 2 ($N = 91$) and the percentages of conclusions that participants drew for themselves

	Second Premise			
	Categorical Premises		Modal Premises	
First Premise	*c is one of those A*	*c is not B*	*c may be one of those A*	*c may not be B*
All those A are B	1. A B c same A B	2. A B same A B ¬ B c	3. A B c A B A B A B ¬ B c	4. A B A B c A B A B ¬B c ¬ B
	∴ c is B: 95	**∴ c is not A: 76**	**∴ c may be B: 72** ***∴ c may or may not be B:* 10**	**∴ c may not be A: 55** ***∴ c may or may not be A:* 14**
None of those A is B	5. A ¬B c same A ¬B	6. A ¬B c A ¬B A ¬B A ¬B ¬B c	7. A ¬B c A ¬B A ¬B A ¬B ¬A B c	8. A ¬B c A ¬B A ¬B A ¬B ¬A B c
	∴ c is not B: 97	***∴ c may be A*: 24** ***∴ c may or may not be not A*: 10** ∴ c is A: 37 **Nothing follows:** 13	**∴ c may not be B: 41** ***∴ c may or may not be B*: 11** ∴ c is not B: 15 Nothing follows: 15	**∴ c may be A: 43** **∴ c may not be A: 17** ***∴ c may or may not be not A:* 10** Nothing follows: 17

Note. Each model is an example from ones that can vary in number and typicality. Results in bold are for necessary conclusions in the model theory, conclusions in Roman font are those that the model theory predicts, whereas those in *italics* are those that the model theory does not predict though they are consistent with it. The balances of percentages in each cell are for other conclusions that participants drew on less than 10% of trials.

5.2. Results

None of the participants failed the checks on their attention. Two independent judges assigned the participant's answers to 12 prior categories, which were the model theory's three predictions and other categories based on the observed sorts of conclusion (Cohen's $k = .99$). Their few disagreements were resolved in discussion.

Table 4 presents the percentages of all the conclusions that participants drew on 10% or more trials for each inference. All eight sorts of inference elicited more of the model theory's predicted conclusions than others (binomial test, $p = .5^8 < .005$). Likewise, 80 participants drew more predicted than unpredicted conclusions, eight drew more unpredicted than predicted conclusions, and there were three ties (binomial test, $p < 1$ in 10 billion; corroborating Prediction 1). Three categorical inferences (1, 2, and 5 in Table 4) yield a necessary conclusion from their intuitive models, whereas Inference 6 and all the modal inferences (3, 4, 7, and 8 do not), and the participants drew more necessary conclusions from the former than from the latter (Wilcoxon test, $z = 4.219$, $p < .001$, effect size: $r = .31$; Prediction 2). Categorical premises tended to yield more necessary conclusions (74%) than modal inferences did (68%), but the difference was not reliable (Wilcoxon test, $z = 1.29$, $p = .2$, effect size: $r = .1$), perhaps because the two instances of Inference 6, with categorical negative syllogistic premises, do not yield a necessary conclusion (see Table 4). Affirmative premises yielded more necessary conclusions (80%) than negative premises did (62%; Wilcoxon test, $z = 6.59$, $p < .001$, effect size: $r = .49$; Prediction 3).

One surprising result was that participants drew 34% of conclusions about possibilities to categorical premises of the sort: *None of those A is a B; c is not one of those B*. They inferred: *c may be B* or an equivalent. It is a necessary conclusion in the model theory, but likely to be a residual effect from other modal inferences in the experiment, because such conclusions did not occur in earlier studies of categorical inferences (Khemlani & Johnson-Laird, 2012, 2022). Another unexpected result, which is also consistent with the model theory, was the use of disjunctive modal conclusions such as *c may or may not be one of those A*. Only 24 participants drew such conclusions, and so, though they follow of necessity, they may be idiosyncratic. Overall, the results bore out the model theory's three main predictions.

6. Experiment 3: Inferences from affirmative and negative premises

In this experiment, the participants drew their own conclusions from pairs of premises, which included affirmative and negative categorical assertions as their first premise, such as

Some of those artists are swimmers,
Some of those artists are not swimmers,

and affirmative and negative assertions about individuals as their second premise, such as

Cynthia is one of those artists,
Cynthia is not one of those artists.

Each of the second premises was either categorical or modal (see Table 5), and so the experiment was a further test of Predictions 1 and 3. Unlike the previous experiment, none of the pairs of categorical premises yields a necessary categorical conclusion, and so Prediction 2 was not under test. However, as we learned from the previous experiment, reasoners sometimes draw conclusions about possibilities from categorical premises. They follow necessarily from these premises, but the computer program for drawing conclusions from categorical premises has no way of producing them (see Khemlani & Johnson-Laird, 2022). As in Experiment 2, the model theory predicts that affirmative premises should elicit more necessary conclusions than negative premises do (Prediction 3).

Table 5
Typical intuitive and deliberative models for the eight pairs of premises in Experiment 3 ($N = 86$), and the percentages of those conclusions that the participants drew on more than 10% of trials for each inference (with necessary conclusions in bold)

	Categorical Second Premises		Modal Second Premises	
	c is A	*c is not B*	*c may be A*	*c may not be B*
Some of those A are B	1. A B c A B A B A B A A c	2. A B A B A B A B A ¬B c A ¬B c	3. A B c A B A B A B A A ¬A c	4. A B A B c A B A B A ¬B c A ¬A B c
	∴ c is B: 45 **∴ c may be B: 38** Nothing follows: 6	**∴ c may be A: 27** ∴ c is not A: 12 **∴ c may not be A: 16** Nothing follows: 15	**∴ c may be B: 62** ∴ c is B: 8 Nothing follows: 10	**∴ c may be A: 22** **∴ c may not be A: 24** ∴ c is A: 5 Nothing follows: 22
Some of those A are not B	5. A ¬B c A ¬B A ¬B A ¬B A A B c	6. A ¬B c A ¬B A ¬B A ¬B A A c	7. A ¬B c A ¬B A ¬B A ¬B A A ¬ A c	8. A ¬B c A ¬B A ¬B A ¬B A A B c
	∴ c is not B: 53 **∴ c may not be B: 17** Nothing follows: 5	∴ c is A: 23 **∴ c may be A: 38** Nothing follows: 10	**∴ c may not be B: 44** **∴ c may be B: 11** ∴ c is B: 0 Nothing follows: 15	**∴ c may be A: 40** ∴ c is A: 11 Nothing follows: 23

Note. Each model is an example of ones that vary in number and typicality. Results in bold are for necessary conclusions in the model theory; the remaining conclusions are those that the model theory predicts. The balances of percentages in each cell are for other conclusions that participants drew on less than 10% of trials.

6.1. Method

6.1.1. Participants

This experiment tested a new sample of 86 psychology undergraduates (75 females, nine males, two non-binaries; mean age 23.4 years, $SD = 9.0$) from the same population as before. The G*power 3.1 program yielded a sample size of 80 in order to detect an effect size of 0.38

of correct inferences with 95% power and an alpha error $p = .05$ using a one-tailed Wilcoxon test.

6.1.2. Design, materials, and procedure

The design manipulated three variables: The first premise was either affirmative (*Some of those A are B*) or negative (*Some of those A are not B*), the second premise either affirmed that an individual, *c*, belonged to *A* (*c is one of those A*) or did not belong to *B* (*c is not one of those B*), and it was either categorical or possible. The resulting eight sorts of pairs of premises are in Table 5, together with their typical intuitive and deliberative models. Each participant carried out two instances of the eight sorts of inference for a total of 16 trials in a different random order. An additional trial was a simple check on whether the participants were paying attention in the experiment. The materials and the procedure were identical to those in Experiment 2, so the participants drew their own conclusions from the pair of premises. The Supporting Information presents the full set of contents in English and in their original Portuguese.

6.2. Results

None of the participants failed the attention check. Two independent judges assigned the participant's conclusions to 12 prior categories as in the previous experiment, and almost always they agreed (Cohen's $k = .99$). They resolved their few disagreements in a discussion. Table 5 presents the percentages of the participants' conclusions occurring on more than 10% of the trials for each sort of inference, for example, disjunctive conclusions, such as *c may or may not be one of those A's*, occurred less often than this criterion for each inference and so do not appear in the table. All eight sorts of inference elicited more of the model theory's predicted conclusions than unpredicted conclusions (binomial test, $p = .5^8 < .005$). Likewise, 72 participants out of 86 drew more predicted conclusions than unpredicted ones, nine participants drew more unpredicted conclusions than predicted ones, and the remaining five participants were tied (binomial test, $p < 1$ in a 100 million; Prediction 1). Because none of the categorical premises has categorical conclusions that follow of necessity, the theory predicts that they will elicit fewer such conclusions (34%) than modal premises elicit (51%; Wilcoxon test, $z = 4.87$, $p < .001$; effect size: $r = .37$). Affirmative premises tended to yield more necessary conclusions (47%) than negative premises did (37%, Wilcoxon test, $z = 4.72$, $p < .001$, effect size: $r = .36$; Prediction 3). No surprising results occurred other than the poorer performance in this study than in the previous ones. Its cause is likely to be the greater number of possible models for the quantifier *some* than for *all* or *none*. Indeed, all the inferences have intuitive models that differ from their deliberative models.

7. Experiment 4: Modal syllogisms

The preceding studies examined inferences based on a quantified premise and a premise about a particular named individual. In contrast, Experiment 4 investigated inferences from two monadic quantified premises, that is, syllogisms, in order to test the model theory's three predictions. A previous study of categorical syllogisms had shown that participants were more likely to accept a given conclusion about a possibility when it followed from an intuitive model of the premises than when it depended on a deliberative model (Evans et al., 1999). But no previous experiment had examined syllogisms with a premise about a possibility, even though they are a major sort in Aristotle's *Prior Analytics* (Barnes, 1984; and for a defense of his account, see Malink, 2013), and psychologists have studied categorical syllogisms for over a century (Störring, 1908).

The following sort of categorical premises

Some of those A are B
All those B are C

has the following typical intuitive model:

A	B	C
A	B	C
A		

which yields the conclusion:

∴ *Some of those A are C.*

And this conclusion is necessary: Deliberation yields no model that refutes it. In contrast, the following modal premises

Some of those A may be B
All those B are C

have the same intuitive model as the one above, but a deliberative model of the negation of the corresponding categorical (*None of those A is a B*) for the first premise:

A	¬ B	
A	¬ B	
	B	C
	B	C

The two models together support a necessary inference of the modal conclusion:

∴ *Some of those A may be C.*

As this contrast illustrates, categorical syllogism that yield necessary categorical conclusions from their intuitive and deliberative models should yield more of them than parallel modal syllogisms.

7.1. Method

7.1.1. Participants

This experiment tested a new sample of 86 students from the same population as before (69 females, 16 males, and one non-binary; mean age 23.2 years, *SD* = 9.0). The G*power 3.1 program yielded a sample size of 80 in the same conditions as Experiment 3.

7.1.2. Design

The experiment investigated eight sorts of syllogism, which were a result of three manipulations. The first premise was either affirmative (*Some of those A are B*) or negative (*Some of those A are not B*), and either categorical or modal (*Some of those A may be B*), and the second premise was either affirmative (*All B are C*) or negative (*No B is C*). Table 6 presents the eight sorts of premises and examples of their typical intuitive and deliberative models. The categorical syllogisms have valid categorical conclusions for only two of the four inferences, that is, those with affirmative first premises (Inferences 1 and 2). The participants drew their own conclusions from two instances of each sort of syllogism with different contents. There were two simple attention checks. Each participant was randomly assigned to one of the two versions of the materials and received the 18 trials in a different random order.

7.1.3. Materials and procedure

The contents for the 16 syllogisms (two instances of the eight sorts of syllogisms) were based on those in the previous experiments, where A named a profession, B named an avocation, and C referred to a characteristic, such as *nice* or *agile*. We assigned the resulting contents twice at random to the set of syllogisms in order to yield two versions of them. The procedure was the same as in the previous experiments.

7.2. Results

None of the participants failed the attention checks. As before, two independent judges assigned the participant's responses to 12 categories and almost always agreed (Cohen's $k = .98$). They resolved disagreements in a discussion. Table 6 presents the percentages of conclusions that the participants drew for themselves from the eight sorts of syllogistic premises. Unpredicted conclusions all occurred on less than 10% of trials with each relevant inference. Hence, all eight sorts of syllogism elicited more predicted conclusions than unpredicted conclusions (binomial test, $p = .5^8 < .005$). Likewise, 71 participants out of 85 drew more predicted conclusions than unpredicted ones, 10 participants drew more unpredicted conclusions than predicted ones, and the remaining four participants were tied (binomial test, $p < .1$ in 100 million; Prediction 1). The predicted conclusions include erroneous categorical conclusions to inferences based on the first premise that refers to a possibility. They were common for inferences based on affirmative premises. For instance, as Table 6 shows, to premises of the sort:

Some of those A may be B
All B are C

participants tended to draw conclusions of the modal sort (41%):

Some of those A may be C.

But an almost equal number drew conclusions of a categorical sort (35%):

Some of those A are C.

Table 6
Typical instances of intuitive and deliberative models for the eight pairs of premises in Experiment 4 ($N = 86$) and the percentages of predicted conclusions that participants drew for themselves (with predicted necessary conclusions in bold)

First Premise	Categorical Second Premise			
	All B are C		*No B is a C*	
Some of those A are B	1. A B C Same		2. A B ¬C A B ¬C	
	A B C		A B ¬C A B ¬C	
	A		A A C	
	∴ Some of those A are C:	**74**	**∴ Some of those A are not C:**	**54**
	Nothing follows:	8	Nothing follows:	13
Some of those A may be B	3. A B C A ¬B		4. A B ¬C A ¬B C	
	A B C A ¬B		A B ¬C A ¬B C	
	A B C		A B ¬C	
	∴ Some of those A may be C:	**41**	**∴ Some of those A may not be C:**	**26**
	∴ Some of those A are C:	35	∴ Some of those A are not C:	22
	Nothing follows:	12	Nothing follows:	15
Some of those A are not B	5. A ¬B A ¬B C		6. A ¬B A ¬B C	
	A A C		A A C	
	B C A B C		B ¬C B ¬C	
	∴ Some of those A are not C:	30	∴ Some of those A are not C:	30
	∴ Some of those A are C:	22	∴ Some of those A are C:	22
	Nothing follows:	**21**	**Nothing follows:**	**25**
Some of those A may not be B	7. A ¬B A B C		8. A ¬B A ¬B C	
	A A B C		A A B ¬C	
	B C A B C		B ¬C A B ¬C	
	∴ Some of those A may not be C:	**24**	**∴ Some of those A may not be C:**	**19**
	∴ Some of those A are C:	20	**∴ Some of those A may be C:**	**11**
	Nothing follows:	24	∴ Some of those A are not C:	14
			Nothing follows:	28

Note. The models are examples with minimal numbers of tokens. The balances of percentages in each cell are for conclusions that occurred on less than 10% of trials for the inference.

The latter is predicted from the nature of the single intuitive model of the premises. It is open to other putative explanations—the participants forgot the modal because other inferences were categorical. Other studies of sentential reasoning (with no quantifiers) have also observed that individuals often draw categorical conclusions to modal premises (Ragni & Johnson-Laird, 2020; Experiment 5).

Necessary conclusions follow from the categorical syllogisms but only from those with an affirmative first premise (Inferences 1 and 2). They elicited a greater percentage of necessary conclusions (64%) than those with a modal first premise (33%; Wilcoxon test, $z < 5.84$, $p < 1$ in 10 million, effect size: $r = .45$; Prediction 2). Discounting the two categorical syllogisms with no necessary conclusions (Inferences 5 and 6), the remaining six corroborated the model theory's prediction that there were more necessary conclusions from syllogisms with an affirmative second premise (46%) than from those with a negative second premise (37%; Wilcoxon test, $z = 3.68$, $p < .0005$; effect size: $r = .28$; Prediction 3). The failure of participants to draw conclusions about the possibilities of categorical syllogisms was surprising because they had done so in the previous study. It may be that syllogisms are sufficiently difficult to inhibit modal conclusions. Low accuracy and high variation in conclusions are typical of studies of syllogistic inference in which individuals draw their own conclusions: These inferences have a greater variety of models than the inferences in the three preceding experiments (see, e.g., Johnson-Laird & Steedman, 1978; Khemlani & Johnson-Laird, 2022). Overall, Experiment 4 bore out the predictions of the model theory, and they corroborated the common error of drawing a categorical conclusion to modal premises.

8. On cognitive theories of human reasoning based on modal logic

A cognitive theory of human reasoning embodying a modal logic should be welcomed because it would perpetuate the Aristotelian tradition that the task of logic is to analyze argumentation. Hence, this section considers the problems and prospects for such a theory of quantified modal logic. It outlines obvious discrepancies between reasoning and logic (Section 8.1), potential logics for the foundations of the theory (Section 8.2), the role of pragmatics in translating the vernacular into the formal language of logic (Section 8.3), and the implications of the evidence from our experiments for such a theory (Section 8.4).

8.1. Prior differences between reasoning and logics

Large differences exist between human reasoning and *standard* logics, which we define as any logic that includes the classical sentential calculus. This calculus has fixed interpretations of negation, *not*, and the sentential connectives, *and, if*, and *or*. Their meanings are idealizations of those in everyday discourse and defined in terms of the truth or falsity of their clauses, for example, a conjunction formed with *and* is true if both its clauses are true and otherwise it is false. Standard logics therefore include quantificational calculi, which concern analogous idealizations of *some* and *all* and higher-order quantifiers such as *few* and *most* and normal modal logics, which add idealizations of *possible* and *necessary* to the sentential calculus or to a predicate calculus (Kripke, 1963). The major differences are as follows:

1. Human reasoners can draw their own conclusions from premises. In contrast, a standard logic yields infinitely many valid conclusions from any premises whatsoever, for example, a conjunction of single premise with itself five times. Hence, it cannot single out any particular conclusion to draw on logical grounds alone. The major theory of reasoning based on the formal rules of a predicate calculus accordingly did not explain how reasoners could draw their own conclusions (see, e.g., Rips, 1994).
2. Human reasoners are unfazed by an inconsistency. In perhaps the only record of interchanges between Wittgenstein and Turing (see Wittgenstein & Bosanquet, 1989, pp. 207 et seq.) they discussed the consequences of the inconsistency of a speaker who asserts the paradox of the liar: "I am lying." Turing remarks that a self-contradiction is usually a criterion of having done something wrong, "but in this case one cannot find anything done wrong." Wittgenstein goes further and argues that one has not done anything wrong. Elsewhere he castigates mathematicians for their superstitious fear of inconsistencies (Wittgenstein, 1964, p. 144). Yet, in standard logic, any conclusions whatsoever follow from inconsistent premises. An inference of this sort is valid:
   The Mona Lisa is in the Louvre.
   The Mona Lisa is not in the Louvre.
   ∴ *There is a hippopotamus in the kitchen.*
   It has no counterexample in which the premises are true and the conclusion is false because no case can occur in which the premises are true. That is why, in part, the model theory replaces validity with necessity, and the inference is not necessary, because it describes a situation that the premises do not (see Section 2.5). Hence, as in everyday life, the effects of inconsistency are local: Reasoners seek to resolve it (Johnson-Laird et al., 2004).
3. Human reasoners treat everyday inferences as defeasible (aka "nonmonotonic" in AI), for example, they are happy to withdraw conclusions. In a standard logic, a conclusion that is false might signal that a valid inference should not be drawn, but it cannot justify the withdrawal of the conclusion. Its direct denial is inconsistent with the premises and justifies any conclusion whatsoever, including the one to be withdrawn.
4. Human reasoners draw different conclusions from the same premises from one time to another, and the two conclusions can be inconsistent with one another, for example, in one study, 95% of participants improved their inferential accuracy in syllogistic reasoning from 1 week to the next, and they had no knowledge of the forthcoming second test (Johnson-Laird & Steedman, 1978). The model theory predicts this difference on the grounds that reasoning is not a deterministic process and that reasoners are more likely to rely on intuition alone the first time than the second time.
5. Human reasoners use a vernacular that contains a variety of truth values, which are not segregated to the semantic system as they are in standard logics (see Johnson-Laird, Byrne, & Khemlani, 2023). Speakers can therefore make paradoxical assertions, such as this version of the "liar":

This assertion is false.

And they can use complex truth values, such as

It is true and it could not have been false.
It could be true and it could be false.
It is neither true nor false.

8.2. Some candidate logics for a cognitive theory

Despite the preceding differences, there is a potential for a two-way influence between modal logics and the psychology of reasoning. A particular modal logic could be the basis of a theory of reasoning in cognitive science (see Bringsjord & Govindarajulu, 2020). Indeed, a reviewer suggested that system KD45 might be such a candidate. It has the plausible consequence that beliefs do not have to include all their logical consequences, but the implausible consequence is that contrary to the model theory, two possibilities that are consistent with one another cannot be condensed into one: Such inferences are invalid (see Meyer & van der Hoek, 1995, p. 173, who considered KD45 for an account of beliefs). And the model theory applies to assertions that their speakers do not believe, such as those that are mere hypotheses or that correspond to other people's beliefs.

One way to cope with the withdrawal of conclusions, as one reviewer suggested, might be to adopt a nonmonotonic modal logic, that is, a modal logic that allows conclusions to be drawn tentatively and later withdrawn. There are many such proposals—from McDermott (1982) to Areces, Cassano, Fervari, and Hoffmann (2023). They also have their potential shortcomings, in particular in combining logical validity with the withdrawal of conclusions (cf. Stalnaker, 1993). The principal accounts are based on standard modal logics, such as system S5, which the reviewer found implausible, to which they add methods to allow nonmonotonocity. The typical systems strike us as implausible psychologically, for example, the truth of "possibly Trump is in Manhattan" depends on a failure to prove that he is not. This method may work well in the management of a database, but to argue, say, that telepathy may work because it has not been proved not to work is too weak to rely on in daily life. Nonetheless, a novel approach to nonmonotonicity in a modal framework might be viable.

Another reviewer remarked that a "contingency" logic in which *possibly A and possibly not-A* is treated as a primitive would predict all our results but gave no example of such a prediction. The reviewer cited a contingency logic due to del Cerro and Herzig (2011). It takes as its primitive operators: affirmative and negative contingency and affirmative and negative necessity. The resulting language is negation-free and seems more pertinent to a programming technique (known as ASP) than to predictions about everyday reasoning.

A third reviewer suggested that we describe how a cognitive theory accommodating an appropriate logic might account for inferential competence, that is, the human ability to make correct inferences. Indeed, it is not easy to show how formal rules of inference can explain systematic errors, such as those that have the force of cognitive illusions (see Section 2.4). Rips (1994) formulated the most successful theory of categorical reasoning based on formal rules for the standard calculus dealing with simple quantifiers, *some* and *all*, but offered no account of systematic errors or of how reasoners formulate conclusions of their own. Likewise, as we stated at the start of the present article, the two most robust phenomena in quantified reasoning are the vast differences in ability from one individual to another and in difficulty from one inference to another. So, we agree with the reviewer's suggestion that the cognitive theory should concern the competence of ideal human reasoners. The reviewer itemized a number of the components of the cognitive theory that can help to accommodate a modal logic, and the next section considers one of them.

8.3. Pragmatics and the recovery of logical form

A major component needed to apply a modal logic to arguments in natural language translates inferences from the vernacular into the symbols of its logical language. In other words, it should extract the logical forms from quantified premises to match those of the formal rules of inference. No algorithm exists to carry out this task. The difficulty, as we have argued many times (e.g., Johnson-Laird, 2006, p. 165 et seq.; Johnson-Laird & Ragni, 2023) is that logical form cannot be derived solely from the syntax of sentences in natural language. It also depends on their meanings. For instance, consider again this inference (from Section 2.3):

Some of the customers ate steak or lobster.
∴ *Some of the customers ate steak.*

It would be a mistake to translate the "or" in the premise into the disjunction of a standard logic, and therefore to reject the inference as invalid. Reasoners treat it as necessary (Rasga et al., 2022).

Pragmatics can help, such as a system based on Grice's (1989) account. He aimed to defend standard logic as underlying everyday discourse. Conversation, he argued, follows certain conventions (or "maxims") that enable speakers to mean more than they say. Disjunctions and other connectives in natural language have the meanings of standard logic, and those who disagree overlook these conventions (ibid., p. 24). One convention is that speakers should not assert less than they know. So, when a speaker says:

Joyce is either in Trieste or Paris,

she implicates that she does not know which of the two cities Joyce is in (ibid. p. 8). Otherwise, she would have named a single city. Conversational implicatures are defeasible, and so they can be canceled without contradiction (ibid. p. 44), as when the speaker adds a coda:

. . . and I know where he is.

Theorists have developed Grice's pragmatics in many ways (e.g., Levinson, 2000; Sperber & Wilson, 1995). A useful method is to use an implicature to create a missing premise, which then allows a valid inference in standard modal logics. Thus, Kratzer and Shimoyama (2002) argued that when a waiter offers you a choice of sole or lobster, he implicates that one choice is possible if and only if the other choice is possible. The addition of this premise to the waiter's remark yields the valid inference that you can have the sole, that is, a valid inference of an *or*-deletion (see Section 2.3). It may be that pragmatics can explain why the earlier example, *Some of the customers ate steak or lobster*, should be interpreted as equivalent to some ate steak and some ate lobster, that is, an "*or*-deletion" inference (see Section 2.3).

In Gricean theory, what elicits conversational implicatures are single utterances as a whole (Cohen, 1971). Yet, as the model theory predicts, *or*-deletions occur in violation of this constraint, and so a Gricean implicature cannot explain the following inference, which depends on three separate assertions:

> Imagine that your professor told you that you are *permitted* to do only one of the following actions:
>
> *You can do your homework.*
> *You can do the presentation slides.*
>
> Are you *permitted* to do your homework?
> Yes ☐ No ☐ Impossible to determine ☐.

Participants tended to respond "yes" (Rasga et al., 2022). However, current pragmatic theories have become highly sophisticated, for example, Bar-Lev and Fox (2020), and so the discovery of optimal pragmatics is one goal for the translation of the vernacular into modal logic.

8.4. The impact of experimental evidence

Several results of our experiments should influence the choice of modal logic on which to build a cognitive theory. We evaluated the status of all the frequent modal inferences in our experiments in various modal logics based on the most elementary one, system K, and others such as system T that are based on it (Kripke, 1963). As we mentioned, a reviewer suggested system KD45 (see Chellas, 1980, p. 132). Logics based on K add to it further axiom schemas and equivalent assumptions about the accessibility of possible worlds, and so inferences that are invalid in K become valid in these logics. Most of these assumptions concern the concatenation of more than one modal operator in the same sentence, such as

> *Possibly it is possible that A is necessary.*

How reasoners in daily life treat such concatenations is largely unknown, though their use often appears to be to emphasize uncertainty, for example, "Perhaps it is possible that Evelyn may be engaged." Hence, for our purposes, the key axiom schemas to add to system K do not concern the concatenation of modals. The first of them is:

If necessarily A then A,

and it converts K into T, the modal logic that Osherson (1976) relied on in his pioneering psychological study. The assumption guarantees that if necessarily it will rain, then indeed it will rain—an inference that is not valid in system K. A second key assumption is:

If necessarily A then possibly A.

It is essential for any plausible deontic logic, such as system D, because if, say, an action is obligatory (deontically necessary), then it is permissible (deontically permissible).

Perhaps the crux is that in standard modal logics, the possibility that individuals do not have a property rules out the possibility that do have that property. For the model theory, however, one of these claims presupposes the other. So, consider a case that we chose arbitrarily:

Some of those bakers are artists.
Possibly, some of the bakers are chefs.
∴ *Possibly, some of those artists are chefs.*

An intuitive model of the premises in which the possibility holds is:

bakers	artists	chefs
bakers	artists	chefs
bakers	artists	
bakers		

The deliberative model is based on the presupposition of the second premise—the possibility that none of the bakers is a chef:

bakers	artists	
bakers	artists	
bakers		
		chef
		chef

The two models together yield the conclusion:

∴ *Possibly, some of those artists are chefs.*

The conclusion is invalid in standard model logics because they treat the possibility of not A as establishing the invalidity of an inference of possibly A, and the deliberative model is such a counterexample.

A corollary is that the model theory implies many more possibilities from premises than standard modal logics do. The model theory allows any possibility that does not contradict a valid inference from the corresponding categorical premises. Perhaps the simplest contrast is between

At least some of the architects are brutalists
∴ *Possibly all the architects are brutalists*

and

At least some of the architects are brutalists
∴ *Possibly none of the brutalists is one of the architects.*

The first inference is necessary in the model theory because its conclusion does not contradict the premise, but the second inference is not necessary because its conclusion does contradict the premise. Both inferences are invalid in standard modal logics.

The goal of a cognitive theory based on a modal logic should be to account for the inferential competence of an ideal reasoner. The discrepancies between the model theory and modal logic that we have discussed may well have a solution. A more difficult discrepancy to resolve is one that only experiments published elsewhere address: They show that participants reject inferences that are valid in standard modal logics but that are not necessary in the model theory (e.g., Ragni & Johnson-Laird, 2020). A quantified example worth examining in future research is an inference of this sort:

Possibly all the flights to London are on time.
∴ *Possibly all the flights to London and to Glasgow are on time.*

The model theory predicts that reasoners will reject this inference: It is not necessary because the premise does not refer to flights to Glasgow. Yet, the inference is valid even in system K. A theory based on modal logic could block the inference on pragmatic grounds, but nothing can undermine its validity.

In summary, a cognitive theory of inferential competence founded on a modal logic might consider the principles above and abandon validity as the criterion of correct reasoning in favor of necessity in the sense we defined earlier. In our view, the use of a modal logic as a basis of human competence in reasoning about possibilities should put to one side the question of the withdrawal of conclusions. The goal should be to pin down the crucial factors of necessary inferences about possibilities that distinguish them from logically valid inferences.

9. General discussion

Our paper may have presented the first set of experiments investigating quantified modal reasoning—a topic that goes back to Aristotle (see his *Prior Analytics* in Barnes, 1984), and that is the subject of standard quantified modal logics (see, e.g., Chellas, 1980; Hughes & Cresswell, 1996). The experiments tested the mental model theory, and their overall results corroborated its three main predictions about the conclusions that individuals should draw for themselves. They tended to infer the three sorts of predicted conclusions: those concerning possibilities, those that are categorical, and those that assert that nothing follows from the premises. The story of the experiments can be told in three inferences.

The first inference is:

1. *All those artists are possibly cyclists.*

   Manuela is one of those artists.

The majority of the participants in Experiment 1 drew a conclusion of this sort, where "∴" stands for "therefore":

∴ Manuela is possibly a cyclist.

Most of the others drew a predicted but erroneous categorical conclusion:

∴ *Manuela is a cyclist.*

The model theory explains these results: An assertion that *A is possible* presupposes that *A is not possible* (see Aristotle's similar claim in the epigraph to the present paper), and reasoners represent what is possible in an intuitive model of this sort:

artist	cyclist	Manuela
artist	cyclist	

where we use words to stand in for a model of actual artists and cyclists, and each row in the diagram depicts a different individual. Reasoners who go no further will treat this model as a single possibility. And when there is only one possibility, it corresponds to a fact, and so these reasoners will draw the categorical conclusion above. But if they deliberate then they can construct a model of the presupposition that it is possible that not all those artists are cyclists:

artist	¬ cyclist	Manuela
artist	cyclist	

where "¬" is a symbol for negation, and so Manuela is not a cyclist in this model. The two models together yield the necessary conclusion: *Manuela is possibly a cyclist*.

The second inference is one in which the modal premise is the one about the individual. We use the same contents as those in the first inference as an aid to readers, but no participant in any of our experiments encountered particular contents more than once:

2. *All those artists are cyclists.*

   Manuela is possibly one of those artists.

An intuitive model is the same sort as before, and the deliberative model is one in which Manuela is possibly not one of those artists:

artist	cyclist	
artist	cyclist	
¬ artist		Manuela

The two models together predict the most frequent conclusion that the participants drew, one that is the same as before:

∴ *Manuela is possibly a cyclist.*

The intuitive model predicts the most frequent error, the corresponding categorical conclusion.

The third inference (from Experiment 3) depends on a premise based on the quantifier, "some," and a modal premise about an individual:

3 *Some of those artists are cyclists.*

Manuela may be one of those artists.

The premises have an intuitive model of the sort:

artist	cyclist	Manuela
artist	cyclist	
artist		

and a deliberative model of the sort:

artist	cyclist	
artist	cyclist	
artist		
¬ artist		Manuela

As the models predict, the most frequent conclusion that the participants drew was of the sort:

∴ *Manuela may be a cyclist,*

and the most frequent error was the corresponding categorical conclusion. An independent study carried out in English (Ragni & Johnson-Laird, 2020; Experiment 4) corroborated erroneous categorical conclusions from premises about possibilities but without quantifiers.

The three inferences seem simple and obvious. And they do not seem to reflect the use of "those artists" in a premise instead of "the artists," which in Portuguese (*os artistas*) has no implication that a relevant set of artists exists (cf. Boolos, 1984; and the reply in Johnson-Laird & Bara, 1984). Readers can ask themselves what conclusions, if any, they would draw from the three inferences. Consider, for instance, this example based on Inference 3): If some of those artists in the 2024 Whitney Biennial are cyclists, and Nikita Gale may be one of those artists, it seems clear that she may also be a cyclist.

The validity of an inference in any particular standard modal logic depends on the logic's assumptions about the relevance (aka "accessibility") of one possible world to the truth values of sentences in another possible world, typically the real world. Of our three inferences above, Inference 1 is valid in standard modal logics, including system K, the simplest one of all, and therefore it is valid in all the modal logics based on it, including system T, which Osherson (1976) used as the basis of his pioneering experiments investigating modal reasoning based on sentential connectives.

Inference 2 is invalid in any modal logic from K to KD45 that has a counterexample in which the first premise, *All those artists are cyclists*, holds in the real world, but the second premise holds in a possible world in which Manuela is one of those artists, but they are not all cyclists. So, one way to transform the inference into a valid one is to make the first premise necessarily the case so it then holds in all possible worlds relevant to the truth or falsity of sentences referring to the real world. This assumption seemed implausible to us, but a reviewer cited an argument due to Chung and Mascarenhas (2023): Speakers treat their own assertions as epistemically necessary. A difficulty for this assumption is that epistemic possibilities in everyday life run from possibility to certainty rather than necessity (Lassiter, 2017; White, 1975). And do speakers always believe what they assert? The claim is doubtful if only because lies would be impossible.

Inference 3 is also invalid in standard modal logics, and as the reviewer allowed, it remains so even if both premises are in addition asserted to be necessary. It is invalid in any modal logic that allows a counterexample in which in the real world, Manuela and two others are artists but only the two others are cyclists, and hence Manuela is not a cyclist, and so it is false that she may be one.

Is there a confound in our results from our use "those artists" (*aqueles artistas* in Portuguese)? In the premises of every inference in our experiments, we used such a phrase, for example:

All those artists are possibly cyclists.

A reviewer argued that as Lewis and others would argue, the word *those* ties the artists to the actual world, whereas the modal operator requires us to consider a possible world for the cyclists, in which only the counterparts to those artists exist. In fact, the late David Lewis wrote about an example: "If I'd asked the boss for a raise, he'd have granted it," depends on the sphere of possible worlds in which the boss was in a generous mood (personal communication to J-L, October 10, 1983). So, it is essential to Lewis's realism about possible worlds that definite descriptions, such as "the boss," refer both to individuals in the real world and to corresponding individuals in possible worlds. However, the essential point is that even if the reviewer's claim was correct, it does not explain our results. A premise of the sort:

All those A's are possibly B's,

in which A's exist in reality but have only counterparts in the possible world of B's should surely inhibit participants from drawing conclusions. But it did not, and they drew them in the three inferences above whether the modal premise was the quantified one (Inference 1) or the one about an individual (Inferences 2 and 3). What did inhibit the participants from drawing modal conclusions was the lack of a term common to both premises (Table 3), or both premises being negative (Tables 4–6), and each of these results corroborated the model theory. In sum, the claim about "those" does not explain our results.

Necessary inferences refer only to possibilities or facts to which the premises refer, and deny none of them. Necessity in this sense replaces logical validity as the model theory's criterion for correct reasoning, and it was the most frequent sort of conclusion to the premises above, and to most inferences that have them in our experiments. There was also a minority of erroneous categorical conclusions of the sort that intuitive models predict, and when participants could not integrate premises into a single model, they tended to respond that nothing follows (see Inference 4 in Table 3). The three sorts of conclusion that the model theory predicts occurred in all four experiments (Prediction 1 in Section 3.3). The theory predicts that necessary conclusions should occur more often from categorical premises than from corresponding ones referring to possibilities, and they did so in the two experiments, Experiments 2 and 4, that examined the difference (Prediction 2 in Section 3.3). The theory also predicts that affirmative premises should yield more necessary conclusions than negative premises, and they did so in the three experiments, Experiments 2–4, that examined the difference (Prediction 3 in Section 3.3).

Could there be some other psychological theory that explains quantified modal reasoning? There are at least a dozen theories of categorical syllogisms. None makes as accurate predictions as the model theory does (see Khemlani & Johnson-Laird, 2012, 2022), and none deals with possibilities. To speculate about how they might accommodate such inferences is not easy, but to give readers the flavor of the exercise, one of these theories combines heuristics with probabilities (Chater & Oaksford, 1999). It might explain conclusions about possibilities—as non-numerical probabilities, but to account for the erroneous categorical conclusions inferred from possibilities, it needs to explain why reasoners would assign them a probability of 100%. We are unable to think of a solution to this problem. Yet the theory's use of heuristics is analogous to those in the mReasoner program, which simulates the model theory of categorical syllogisms (see Section 3.1). The model theory's heuristics are akin to the "atmosphere" effect in syllogistic reasoning, for example, the occurrence of "some" in the premises predisposes reasoners to draw a conclusion containing "some" (e.g., Begg & Denny, 1969; Woodworth & Sells, 1935). So, too, if at least one premise is about a possibility, an acceptable conclusion should be too. But atmosphere fails to explain why the participants' most frequent errors were categorical conclusions from premises about possibilities.

An ideal basis for a theory of quantified modal reasoning would be a modal logic, and we have considered in some detail how such a theory of human competence in reasoning might be realized. Our analysis in Section 8 illustrated some of the problems that would need to be solved, such as the withdrawal of valid conclusions, the infinite consequences of inconsistency, and the prediction of the conclusions that reasoners draw for themselves. The principal discrepancy is between the model theory's account of epistemic possibilities—those based on empirical knowledge of the world—and the concept of possibility underlying standard modal logics. The model theory treats the possibility of an event as presupposing the possibility of its non-occurrence. In standard modal logics, the two do not mutually imply one another. The reason is that in system T and all the logics it underlies, an inference from a categorical assertion of an event implies its possibility, for example:

It is hot.
∴ *Possibly it is hot.*

So, if this conclusion were to imply:

∴ *Possibly it is not hot*,

then this inference would also follow:

It is hot.
∴ *Possibly it is not hot*.

And this conclusion contradicts the premise. Hence, in standard modal logics, if there is a counterexample to a categorical inference, then it is also a counterexample to the corresponding inference about a possibility. For example, an inference such as

Every day it is hot or it rains, or both.
∴ *Every day it is hot*.

has a counterexample: *Every day it is not hot but it rains*. It establishes the truth of the premise and the falsity of the conclusion. Likewise, the corresponding modal inference:

It is possible that every day it is hot or it is raining,
∴ *It is possible that every day it is hot,*

has the same counterexample in standard modal logics. Yet, the inference is necessary in the model theory, and individuals are likely to accept it (cf. Hinterecker et al., 2016).

The model theory's approach also applies to counterfactual possibilities—those that were once possible but that did not happen (e.g., Espino, Byrne, & Johnson-Laird, 2020; Quelhas, Rasga, & Johnson-Laird, 2018). In a situation in which a person knocked a vase off a table but was able to deflect its fall, consider the following counterfactual claim:

The vase would otherwise have hit the concrete floor, not the carpet.

To verify the assertion, you need a kinematic model to simulate the trajectory of the vase had it not been deflected. Umpires in many sports are called on to make such simulations, and in those cases in which the Hawkeye TV system makes them too, it shows that the humans are quite accurate in assessing the truth or falsity of counterfactual possibilities (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2021; Johnson-Laird et al., 2023).

Many complications occur in quantified modal reasoning that the model theory has yet to address, and we itemize the most important ones:

1. Complex quantifiers based on restrictive relative clauses, such as *Most of those who know some of the sculptors.*
2. Multiply-quantified assertions, such as *Everybody admires anyone who admires someone* (for categorical studies, see Cherubini & Johnson-Laird, 2004; Johnson-Laird, Byrne, & Tabossi, 1989).
3. Quantifiers within the scope of modal operators leave open whether an individual exists, such as *Possibly, there is someone that Evelyn wants to marry* (see, e.g., Girle, 2009, Ch. 8; Quine, 1961).
4. Multiple modal operators in a single sentence, such as *It is* highly possible *that the storm* may perhaps *become a hurricane.*

Our study was an initial foray into quantified modal reasoning. It bore out a reasonable hypothesis: a possible event needs an intuitive model of its occurrence and a deliberative model of its presupposed non-occurrence. The two together warrant a necessary inference of a possibility. Human reasoners draw inferences of this sort, which seem to be beyond the explanatory power of standard modal logics. The theory calls for further experimental tests. It may be that a modal logic outside the scope of the standard varieties could account for human competence. And it may be that an alternative but yet to be devised account overturns the model theory.

Acknowledgments

We thank Ana Lapa for helping to develop the materials and for carrying out the experiments. We are also grateful to the following colleagues for help and advice: Maya Bar-Hillel, Ruth Byrne, Sangeet Khemlani, Philipp Koralus, Salvador Mascarenhas, Marco Ragni, Mark Steedman, and Peter Fonagy, who made possible a meeting in the Department of Psychology, University College London, July 20–21, 2023, on the present and other related topics. And we thank three anonymous reviewers for their insightful critiques and Rick Dale for his help and advice in formulating a publishable paper. The first author acknowledges the contribution of the APPsyCI (Applied Psychology Research Center–Capabilities and Inclusion) FCT/UIDB/05299/2020.

References

Areces, C., Cassano, V., Fervari, R., & Hoffmann, G. (2023). DefTab: A tableaux system for sceptical consequence in default modal logics. In R. Ramanayake & J. Urban (Eds.), *International conference on automated reasoning with analytic tableaux and related methods* (pp. 37–48). Cham, Switzerland: Springer Nature.

Bar-Lev, M. E., & Fox, D. (2020). Free choice, simplification, and innocent inclusion. *Natural Language Semantics*, *28*, 175–223.

Barnes, J. (Ed.). (1984). *The complete works of Aristotle* (Vol. 1). Princeton, NJ: Princeton University Press.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences*, *22*, 637–660.

Begg, I., & Denny, J. (1969). Empirical reconciliation of atmosphere and conversion interpretations of syllogistic reasoning. *Journal of Experimental Psychology*, *81*, 351–354.

Boole, G. (1847). *The mathematical analysis of logic*. Cambridge: Macmillan, Barclay, & Macmillan.

Boolos, G. (1984). On 'syllogistic inference'. *Cognition*, *17*, 181–182.

Bringsjord, S., & Govindarajulu, N. S. (2020). Rectifying the mischaracterization of logic by mental model theorists. *Cognitive Science*, *44*, e12898.

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, *23*, 247–303.

Bucciarelli, M., & Johnson-Laird, P. N. (2024). Bruno Bara, mental models, and syllogisms. *Sistemi Intelligenti, A special issue in memory of Bruno Bara*, in press.

Byrne, R. M. J. (2005). *The rational imagination: How people create alternatives to reality*. Cambridge, MA: MIT Press.

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, *38*, 191–258.

Chellas, B. F. (1980). *Modal logic: An introduction*. New York: Cambridge University Press.

Cherubini, P., & Johnson-Laird, P. N. (2004). Does everyone love everyone? The psychology of iterative reasoning. *Thinking & Reasoning*, *10*, 31–53.

Chung, W., & Mascarenhas, S. (2023). Modality, expected utility, and hypothesis testing. *Synthese*, *202*, 10–40.

Cohen, L. J. (1971). Some remarks on Grice's views about the logical particles of natural language. In Y. Bar-Hillel (Ed.), *Pragmatics of natural language* (pp. 50–68). Dordrecht, The Netherlands: Reidel.

Craik, K. (1943). *The nature of explanation*. Cambridge: Cambridge University Press.

del Cerro, L. F., & Herzig, A. (2011). Logics of contingency. *AAAI Spring Symposium Series (SS-11-06): Logical Formalizations of Commonsense Reasoning*, Stanford, CA (pp. 119–122).

de Neys, W. (2022). Advancing theorizing about fast-and-slow thinking. *Behavioral and Brain Sciences*, *46*, e111.

Espino, O., Byrne, R. M. J., & Johnson-Laird, P. N. (2020). Possibilities and the parallel meanings of factual and counterfactual conditionals. *Memory & Cognition*, *48*, 1263–1280.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278.

Evans, J. S. B. T., Handley, S. J., Harper, C. N. J., & Johnson-Laird, P. N. (1999). Reasoning about necessity and possibility: A test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1495–1513.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Feiman, R., Mody, S., & Carey, S. (2022). The development of reasoning by exclusion in infancy. *Cognitive Psychology*, *135*, 101473.

García-Madruga, J. A., Moreno, S., Carriedo, N., Gutiérrez, F., & Johnson-Laird, P. N. (2001). Are conjunctive inferences easier than disjunctive inferences? A comparison of rules and models. *Quarterly Journal of Experimental Psychology*, *54A*, 613–632.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*, *128*, 936–975.

Girle, R. (2009). *Modal logics and philosophy* (2nd ed.). New York: Routledge.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 586–596.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Hinterecker, T., Knauff, M., & Johnson-Laird, P. N. (2016). Modality, probability, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1606–1620.

Hughes, G. E., & Cresswell, M. J. (1996). *A new introduction to modal logic*. London: Routledge.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. London: Routledge & Kegan Paul.

Jeffrey, R. (1981). *Formal logic: Its scope and limits* (2nd ed). New York: McGraw-Hill.

Johnson-Laird, P. N. (1975). Models of deduction. In R. Falmagne (Ed.), *Reasoning: Representation and process* (pp. 7–54). Springdale, NJ: Erlbaum.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.

Johnson-Laird, P. N. (2006). *How we reason*. New York: Oxford University Press.

Johnson-Laird, P. N., & Bara, B. G. (1984). Logical expertise as a cause of error. *Cognition*, *17*, 183–184.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N., Byrne, R. M. J., & Khemlani, S. (2023). Human verifications: Computable with truth values outside logic. *Proceedings of the National Academy of Sciences*, *120*(40), e2310488120.

Johnson-Laird, P. N., Byrne, R. M. J., & Tabossi, P. (1989). Reasoning by model: The case of multiple quantification. *Psychological Review*, *96*, 658–673.

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review*, *111*, 640–661.

Johnson-Laird, P. N., Khemlani, S. S., & Goodwin, G.P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, *19*, 201–214.

Johnson-Laird, P. N., Quelhas, A. C., & Rasga, C. (2021). The mental model theory of free choice permissions and paradoxical disjunctive inferences. *Journal of Cognitive Psychology*, *33*(8), 951–973.

Johnson-Laird, P. N., & Ragni, M. (2019). Possibilities as the foundation of reasoning. *Cognition*, *193*, 130950.

Johnson-Laird, P. N., & Ragni, M. (2023). *Reasoning about possibilities: Modal logic, possible worlds, and mental models.* Under submission.

Johnson-Laird, P. N., & Steedman, M. J. (1978). The psychology of syllogisms. *Cognitive Psychology*, *10*, 64–99.

Johnson-Laird, P. N., & Wason, P. C. (1970). A theoretical analysis of insight into a reasoning task. *Cognitive Psychology*, *1*, 134–148.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Strauss, Giroux.

Kamp, H. (1973). Free choice permission. *Proceedings of the Aristotelian Society*, *74*, 57–74.

Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science*, *42*, 1887–1924.

Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, *138*, 427–457.

Khemlani, S. S., & Johnson-Laird, P. N. (2017). Illusions in reasoning. *Minds and Machines*, *27*, 11–35.

Khemlani, S., & Johnson-Laird, P. N. (2022). Reasoning about properties: A computational theory. *Psychological Review*, *129*, 289–312.

Khemlani, S. S., Mackiewicz, R., Bucciarelli, M., & Johnson-Laird, P. N. (2013). Kinematic mental simulations in abduction and deduction. *Proceedings of the National Academy of Sciences*, *110*, 16766–16771.

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2014). The negations of conjunctions, conditionals, and disjunctions. *Acta Psychologica*, *151*, 1–7.

Kratzer, A., & Shimoyama, J. (2002). Indeterminate pronouns: The view from Japanese. In Y. Otsu (Ed.), *Third Tokyo conference on psycholinguistics* (pp. 1–25). Tokyo: Hituzi Syobo.

Kripke, S. (1963). Semantical considerations on modal logic. *Acta Philosophica Fennica*, *16*, 83–94.

Kyburg, H. -E., Jr. (1983). The role of logic in reason, inference, and decision. *Behavioral and Brain Sciences*, *6*, 263–273.

Lassiter, D. (2017). *Graded modality: Qualitative and quantitative perspectives*. New York: Oxford University Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge: MIT Press.

Malink, M. (2013). *Aristotle's modal syllogistic*. Cambridge, MA: Harvard University Press.

Marek, W., & Truszynski, M. (2013). *Nonmonotonic logic*. New York: Springer.

Markovits, H. (2023). Fast logic and belief-bias: It's less how smart you are than how you think. *Journal of Cognitive Psychology*, *35*(2), 248–254.

McDermott, D. (1982). Nonmonotonic logic II: Nonmonotonic modal theories. *Journal of the Association of Computing Machinery*, *29*, 33–57.

Meyer, J. -J., & van der Hoek, W. (1995). *Epistemic logic for AI and computer science*. Cambridge: Cambridge University Press.

Montague, R. (1974). *Formal philosophy: Selected papers*. New Haven, CT: Yale University Press.

Newstead, S. E., & Griggs, R. A. (1983). Drawing inferences from quantified statements: A study of the square of opposition. *Journal of Verbal Learning and Verbal Behavior*, *22*, 535–546.

Oakhill, J. V., & Johnson-Laird, P. N. (1985). The effects of belief on the spontaneous production of syllogistic conclusions. *Quarterly Journal of Experimental Psychology*, *37A*, 553–569.

Osherson, D. N. (1976). *Logical abilities in children, Volume 4: Reasoning and concepts*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Piéraut-Le Bonniec, G. (1980). *The development of modal reasoning: Genesis of necessity and possibility notions*. New York: Academic Press.

Priest, G. (2008). *An introduction to non-classical logic* (2nd ed.). Cambridge: Cambridge University Press.

Quelhas, A. C., & Johnson-Laird, P. N. (2017). The modulation of disjunctive assertions. *Quarterly Journal of Experimental Psychology*, *70*, 703–717.

Quelhas, A. C., Rasga, C., & Johnson-Laird, P. N. (2018). The relation between factual and counterfactual conditionals. *Cognitive Science*, *42*, 2205–2228.

Quine, W. V. O. (1961). Reference and modality. In W. V. O. Quine (Ed.), *From a logical point of view* (2nd ed., pp. 139–159). New York: Harper & Row.

Ragni, M., & Johnson-Laird, P. N. (2020). Reasoning about epistemic possibilities. *Acta Psychologica*, *208*, 103081.

Rasga, C., Quelhas, A. C., & Johnson-Laird, P. N. (2022). An explanation of *or*-deletions and other paradoxical disjunctive inferences. *Journal of Cognitive Psychology*, *34*, 1032–1051.

Rips, L. J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.

Routley, R. (1982). On what there is not. *Philosophy and Phenomenological Research*, *43*, 151–178.

Shtulman, A., & Phillips, J. (2018). Differentiating "could" from "should": Developmental changes in modal cognition. *Journal of Experimental Child Psychology*, *165*, 161–182.

Skovgaard-Olsen, N., Collins, P., & Klauer, K. C. (2023). Possible worlds truth table task. *Cognition*, *238*, 105507.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford, England: Wiley-Blackwell.

Stalnaker, R. (1993). A note on non-monotonic modal logic. *Artificial Intelligence*, *64*, 183–196.

Störring, G. (1908). Experimentelle Untersuchungen über einfache Schlussprozesse. *Archiv für die gesamte Psychologie*, *11*, 1–127.

Vance, J., & Oaksford, M. (2021). Explaining the implicit negations effect in conditional inference: Experience, probabilities, and contrast sets. *Journal of Experimental Psychology: General*, *150*, 354–384.

Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, *11*, 92–107.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140.

White, A. R. (1975). *Modal thinking*. Ithaca, NY: Cornell University Press.

Wittgenstein, L. (1964). *Remarks on the foundations of mathematics.* Oxford, England: Basil Blackwell.

Wittgenstein, L., & Bosanquet, R. G. (1989). *Wittgenstein's lectures on the foundations of mathematics, Cambridge, 1939*. Chicago, IL: University of Chicago Press.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in formal reasoning. *Journal of Experimental Psychology*, *18*, 451–460.

Yang, Y., & Johnson-Laird, P. N. (2000). Illusory inferences in quantified reasoning: How to make the impossible seem possible, and vice versa. *Memory & Cognition*, *28*, 452–465.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information for: Quantified modal reasoning