

Note: In Press at the *Journal of Cognitive Psychology*

Negation: A Theory of its Meaning, Representation, and Use

Sangeet Khemlani	Isabel Orenes	P. N. Johnson-Laird
Princeton University	University de la Laguna	Princeton University

Running title: Negation

December 30th 2011

Word count (main text): 10,293

Please address correspondence to:

Sangeet Khemlani

Intelligent Systems Section, Code 5515

Navy Center for Applied Research in Artificial Intelligence

Naval Research Laboratory

4555 Overlook Ave. SW

Washington, DC 20375

Abstract

This article presents a model-based theory of what negation means, how it is mentally represented, and how it is understood. The theory postulates that negation takes a single argument that refers to a set of possibilities and returns the complement of that set.

Individuals therefore tend to assign a small scope to negation in order to minimize the number of models of possibilities that they have to consider. Individuals untrained in logic do not know the possibilities corresponding to the negation of compound assertions formed with *if*, *or*, and *and*, and have to infer the possibilities one by one. It follows that negations are easier to understand, and to formulate, when individuals already have in mind the possibilities to be negated. The paper shows that the evidence, including the results of recent studies, corroborates the theory.

Keywords: negation, falsity, mental models, comprehension, representation, reasoning, pragmatics.

[Authors' affiliations and acknowledgements]

Sangeet Khemlani, Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, Washington, DC; Isabel Orenes, Departamento de Psicología Cognitiva, Universidad de la Laguna, Tenerife, Spain; Philip N. Johnson-Laird, Department of Psychology, Princeton University, Princeton, NJ.

This research was supported by a National Science Foundation Graduate Research Fellowship to the first author, and by National Science Foundation Grant No. SES 0844851 to the third author to study deductive and probabilistic reasoning. We thank Jay Atlas, Jeremy Boyd, Herb Clark, Alan Garnham, Sam Glucksberg, Adele Goldberg, Geoff Goodwin, Jennifer Heil, Olivia Kang, Philipp Koralus, Mark Liberman, Max Lotstein, Anna Liu, Paula Rubio, Carlos Santamaría, and Elizabeth Sucuyan for their helpful suggestions and criticisms.

Correspondence concerning this article should be addressed to Sangeet Khemlani, Intelligent Systems Section, Code 5515, Navy Center for Applied Research in Artificial Intelligence, Naval Research Laboratory, 4555 Overlook Ave. SW, Washington, DC 20375. E-mail: skhemlani@gmail.com.

Negation is part of all natural languages, yet its psychology is mysterious given that languages contain terms for *true* and *false*. Negation appears to be redundant, because an assertion such as:

1a. Pat is not here

can be paraphrased as:

1b. It is false that Pat is here.

Paraphrases of this sort sometimes lead individuals to confuse falsity with negation. But, it is clear that the two are distinct because a negative assertion can be true or false, and a true assertion can be affirmative or negative. Indeed, an early psychological discovery was of an interaction between the affirmative or negative *polarity* of sentences and the *truth values* of the propositions that they express.

The early discovery of the interaction between polarity and truth value was one of three main results due to Wason and his colleagues (see Wason & Johnson-Laird, 1972, Ch. 2). The first result was that negative assertions are harder to verify than affirmative assertions (Wason, 1959, 1961). The second result concerns the difference between *complementary* predicates, such as “odd” and “even”, which have just one alternative, and *contrastive* predicates, such as “red”, “green”, “blue” etc, which have many alternatives. In negative assertions, individuals often replace the negation of a complementary predicate, such as *not even*, with an affirmation of its complement, *odd*. The third result was that contrastive predicates yield the interaction between polarity and truth value. When individuals verify assertions containing these predicates, they evaluate affirmatives as true faster than as false, but they evaluate negatives as false faster than as

true (Wason & Jones, 1963). Subsequent studies corroborated this interaction (e.g., Gough, 1965; Slobin, 1966). Wason's research showed that any explanation of negation had to take into account both its grammar and meaning (cf. Klima, 1964). It led to the development of psycholinguistic models of the verification of assertions, which included both these factors (see, e.g., Carpenter & Just, 1975; Clark, 1974; Clark & Chase, 1972; Dale & Duran, 2011; Hunt & MacLeod, 1978; Trabasso, Rollins, & Shaughnessy, 1971; Wason & Johnson-Laird, 1972). Wason and his colleagues went on to examine the sorts of stimuli that it is plausible to describe using negation (Wason, 1965).

With hindsight, what is striking about the early studies of negation was their narrow purview. They focused on assertions presented in isolation from any linguistic context, and on tasks concerning the truth or falsity of single clause sentences, such as *The triangle is not above the star*. And they neglected many other ways in which negation occurs, many other tasks such as the formulation of negative assertions, and many other morphemes apart from *not* and *no* that express negation (Dahl, 1979; Horn, 2001). There was accordingly no general theory of the meaning of negation, its mental representation, or its comprehension. The aim of the present article is to propose such a theory rooted in mental models, to derive some novel predictions from it, and to outline the evidence corroborating these predictions.

The article begins with a description of the fundamentals of the theory of mental models – the “model theory” for short. It extends the theory to explain the parsing of negative sentences, the process of understanding them, the resulting mental representations, and the contextual factors that make them easier to understand. This account leads to five principal predictions about negation. The article then reviews the

empirical evidence, which tends to corroborate these predictions. Finally, it draws some general conclusions about negation.

The model theory of negation

A linguistic theory of negation needs to explain the syntax, semantics, and pragmatics of negation; a psychological theory needs to explain how the process of comprehension takes grammar, meaning, and knowledge into account, what is mentally represented as a result of comprehension, and the contextual factors that affect these processes. We outline such a theory based on mental models.

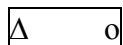
We have described the general theory of mental models in detail elsewhere (e.g., Johnson-Laird, 2006), and so begin with only a brief outline of its main ideas. A major function of language is to enable you to have another person's experience by proxy, and so you construct a mental model of a situation based on the person's remarks. Consider an assertion that a speaker uses to describe the arrangement of two shapes on a blackboard:

2. The circle is on the right of the triangle.

The theory postulates two interleaved processes that occur as you understand the sentence. The first process composes the meaning of constituents (their *intensions*) out of the meanings of their words and the grammatical relations amongst them. Knowledge may modulate the composition, and also help you to determine what is referred to in the assertion. Knowledge suggests, for example, that *on the right of* in (2) has a deictic interpretation concerning the speaker's point of view, because circles and triangles – unlike, say, people and chairs – do not have intrinsic right-hand sides (Miller & Johnson-

Laird, 1976, Sec. 6.1.3). The second process uses intensions to construct or update a mental model of the situation (a representation of the *extension* of the assertion). A mental model represents what is common to a set of possibilities in terms of entities, their properties, and the relations amongst them (Barwise, 1993). Each distinct possibility – in terms of these entities, properties, and relations – requires in principle a separate mental model. Subsequent assertions call for the updating or checking of the models of the description so far, and the role of intensions is crucial for ensuring that any new model remains a model of the assertions in a description. Inference is based on models, and so a deductively valid conclusion is one that holds for all the correct models of the premises. When premises yield multiple models, there is an increase in the processing load on working memory, and, as a result, inferences become more difficult – they take longer and are more prone to error. The logical properties of relational terms emerge from the construction of models. Hence, there is no need for either meaning postulates or for establishing the logical forms of propositions expressed in the use of sentences (Goodwin & Johnson-Laird, 2005).

When you envisage the situation described in the assertion, *the circle is on the right of the triangle*, your mental model makes explicit the relation between the entities, and so it represents a two-dimensional layout isomorphic to the one in this diagram:



The diagram denotes a model with Cartesian coordinates. Its left-to-right axis corresponds to the left-to-right axis of the speaker's point of view of the scene. The mental model is therefore *iconic* in Peirce's sense (see Peirce 1931-1958, Vol. 4, para 418 et seq.): its structure corresponds to the structure of the world under description. It

may also be a representation based on a sensory image (Barsalou, 1999; Glenberg, Robertson, Jansen, & Glenberg, 1999), but images could impede the process of reasoning (Knauff & Johnson-Laird, 2002). Iconic models can, of course, be abstract, e.g., an abstract list-structure can be an iconic representation of one set included within another. In general, mental models based on discourse are as iconic as possible (Johnson-Laird, 2006, Ch. 2).

Mental models are based on a principle of *truth* (e.g., Johnson-Laird & Byrne, 2002). They represent what is true, but not what is false, unless an assertion explicitly refers to falsity. This bias reduces the processing load on working memory. The representation of an exclusive disjunction, such as (3):

3. Either you have the soup or else you don't have the bread

has the following two mental models of what you have, shown here on separate lines, and where “ \neg ” is a symbol for negation:

soup

\neg bread

Each mental model represents a possibility, but the first model does not make explicit that in this possibility it is false you don't have bread, i.e., you *do* have bread, and the second model does not make explicit that in this possibility it is false that you have soup, i.e., you don't have soup. Individuals tend to rely on mental models, and as a result they succumb to systematic fallacies from premises for which valid reasoning depends on what is false (e.g., Khemlani & Johnson-Laird, 2009). But, in some simple tasks, such as enumerating possibilities, individuals can construct *fully explicit* models, which represent both what is true and what is false:

soup bread

\neg soup \neg bread

The contrast between mental models and fully explicit models is relevant, as we show, to the interpretation of negation.

In what follows, we augment the general theory to deal with negation. And here we summarize what is to come. We begin with grammar, because negation is one of its elements. But, negation has an important effect on meaning. The model theory, as we have already described, postulates that individuals use the meaning of an assertion to envisage the possibilities to which it refers: each distinct possibility is mentally represented in an iconic model of what is common to the different ways in which the possibility can occur. The theory postulates that negation refers to the complement of those models to which the corresponding affirmative assertion, or corresponding affirmative constituent, refers. When individuals construct these models, the task can be sufficiently difficult that they have to enumerate the various possibilities one at a time. Such an enumeration is unavoidable for the negation of a sentence containing multiple clauses. The reason is that naive individuals – those who have had no training in logic – do not know the possibilities corresponding to the negation of, say, a conjunction, such as: “It is not the case both that the election is next month and that Viv has registered to vote”. They need to work them out one by one. According to the model theory, the more possibilities that they have to enumerate, the greater the difficulty of the task. But, context can play a role: they should find it easier to construct the complement of a set of models if they have previously represented this set of models. We now spell out the ideas in this summary in more detail.

Negation calls for particular grammatical structures in different languages.

English, unlike many other Indo-European languages, calls for negation in a verb phrase to be preceded by an auxiliary verb that carries tense instead of the main verb. Hence, the negation of:

4a. He laughed.

is:

4b. He didn't laugh.

An important factor is the *scope* of a negation, i.e., what constituents in the intension of the sentence – the proposition that it expresses – are negated. (In future, where the distinction between sentence and proposition is not important, we elide the difference using the term, “assertion”.) Scope is often taken to be a matter of the *logical form* of sentences, but the model theory makes use only of grammatical form. Logical form can be computed only after meaning and context have fixed the proposition that a sentence expresses. But, once this proposition has been determined, sometimes from mental models of the clauses in sentences, logical form is redundant. Indeed, it serves no function in the model theory: individuals can reason instead from the models themselves. Moreover, the complexities that logical form must take into account are quite staggering, e.g., whether or not a relation is transitive depends on the tense of assertions (Goodwin & Johnson-Laird, 2008). No good reason exists to imply that logical form is psychologically real, and this view has been presaged in philosophical analyses of negation (e.g., Atlas, 1977). No existing algorithm can compute the logical forms of all the propositions expressible in natural language, and even some logicians are wary about extending the notion from logic to language. As the late Jon Barwise (1989, p. 4)

remarked: “ ... I find the notion [of logical form] unilluminating. Within the model-theoretic tradition, valid entailments are valid not in virtue of form, but in virtue of content.”

The procedures that the model theory uses to construct a representation of the meaning of an assertion need to identify the argument to which negation applies. According to the theory, this process of identification is biased towards a small grammatical scope rather than a large one, e.g., a verb phrase rather than a sentence. In the sentence: *Not all the critics are kind*, the scope of the negation is the entire clause: *all the critics are kind*. Negation accordingly has a large scope in this case. In contrast, in a sentence such as: *Some of the critics are not kind*, negation has a small scope, just the verb phrase: *are kind*, and it does not include the noun phrase, *some of the critics*. When negation occurs in a verb phrase, it usually has a small scope, but it can have a large scope. This point is clear in a famous example (5a):

5a. The present king of France is not bald.

Its small scope interpretation takes for granted that there is a present king of France, and denies that he is bald. But, its large scope interpretation as paraphrased in:

5b. It is not the case that the present King of France is bald.

does not take for granted the existence of a present king of France (Russell, 1905), and neither does (5c):

5c. The present King of France is not bald, because there is no present King of France.

The scope of “not” is otherwise its siblings and their descendants in the sentence’s parse tree, i.e., whatever it *c-commands* (see, e.g., Chomsky, 1995; Haegeman, 1995, for this syntactic relation). Hence, (6a):

6a. The fire may not burn

can be paraphrased as:

6b. The fire is such that possibly it is not the case that it burns.

The model theory postulates that there is a processing heuristic for negation to be interpreted, where possible, as having a small scope. The principal reason is that a small scope interpretation reduces the number of models that individuals have to represent.

Consider, for instance, how you might respond to the following question:

7. The students are not male adults.

So, who are they?

If you inferred that they are women (female adults), or if you inferred that they are boys (male children), or if you inferred that they are girls (female children), then you assigned a small scope to the negation, that is, you applied it to the individual predicates rather than to their conjunction. In fact, the assertion is consistent with all three possibilities: women, boys, and girls. But, a small scope interpretation reduces scope from:

Not (male and adult)

to:

Not (male) and adult = women

or to:

Male and not(adult) = boys

or to:

Not(male) and not(adult) = girls.

In this way, small scope reduces the number of models of possibilities that individuals need to represent from three to one. As we mentioned in our outline of the model theory, interpretation and inference rapidly increase in difficulty with an increase in the number of possibilities that individuals have to represent (see, e.g., Johnson-Laird, 2006). Hence, it pays to make a small scope interpretation of negation, because it reduces the number of models to be held in working memory.

Another potential reason exists to bias interpretation towards a small scope. It appears to occur for other “sentential operators”, i.e., expressions that can apply grammatically to an entire clause or sentence. These operators include adverbs, such as *surprisingly*, *possibly*, *probably*, and *truly*, as well as phrases based on them. Small scope may reduce the number of models, but it also reduces the complexity of the arguments that an operator takes, and accordingly simplifies computations (see Birney & Halford, 2002; Goodwin & Johnson-Laird, 2005, 2006; Halford, Wilson, & Phillips, 1998). Of course, intonation contour can yield unequivocal interpretations. If a speaker asserts: “the students are not *male* adults”, then the stressed item is the argument of negation, and the remark should elicit the first of the small-scope interpretations above (see, e.g., Bock & Mazella, 1983; Carlson, Frazier, & Clifton, 2009; Moxey, Sanford, & Dawydiak, 2001). Intonational focus in a denial acts as a way to associate the interpretation of the focused element of a sentence with negation, and almost always serves to reduce the scope of negation.

In the model theory, the scope of negation is captured in the representation of the meaning of an assertion, i.e., an intensional representation. There are many ways in

which scope could be represented, and no obvious empirical method to determine which way the mind uses. In mReasoner v. 0.8, our current computational implementation of the theory (Khemlani, Lotstein, & Johnson-Laird, under review), the difference between a large scope and small scope interpretation is represented simply. For sentential negation, the polarity of a sentence or a clause as a whole is set to represent negation. For example, the intensional representation of the assertion:

8. No artists are beekeepers

includes the following information:

Polarity: NIL

The setting of polarity to nil represents that the assertion as a whole is negative. In contrast, the small scope interpretation of:

9. Some artists are not beekeepers

has an intensional representation in which the polarity of the assertion is affirmative but the relation of the set-theoretic inclusion is negative. This implementation works for a wide variety of different sorts of assertion, and it can be expanded to cope with sentences containing several quantifiers.

At the foundation of the theory is its account of the meaning of negation. It postulates that negation is a function that takes a single argument, determined by scope, which refers to a set of models. The *core* meaning of negation is a function that returns the complement of the set. We will explain the implications of this claim, starting with simple *atomic* sentences, i.e., those that do not contain either sentential connectives or quantifiers, such as:

10. The circle is not on the right of the triangle.

As we showed earlier, the mental model of the unnegated sentence is:



The complement of this model calls for a symbol that represents negation:



This use of negation is wholly compatible with Peirce's diagrammatic system of reasoning in which squares representing negation enclose diagrams of conjunctive states of affairs (see, e.g., Johnson-Laird, 2002). The model-building process constructs the model of the corresponding affirmative assertion and applies the negation symbol to the model. Hence, the earlier affirmative assertion and its negation are complements of one another.

Much of language cannot have a perceptual representation, such as a visual image, and negation is one of the most important of such concepts. You could superimpose a visual image of a large red cross on an image of the circle on the right of the triangle (cf. Wittgenstein, 1953). But, you would have to know that the large red cross symbolizes negation, and nothing in the image itself could tell you that. Likewise, you would have to know what negation means, and nothing in an image can capture this meaning – in this case, that in the situation under description the relation in the negated model is false.

In principle, you could construct an iconic representation of each member of the complementary set, that is, of all the positive possibilities compatible with (10), e.g., the circle on the left of the triangle, above it, below it, in front of it, behind it, and so on. But, you need to think of this set of possibilities as an exhaustive disjunction, and nothing in a set of images can represent either that the set is exhaustive or that the images represent a

disjunction of possibilities. It follows that negation has to be represented by a *symbol* (see Peirce, 1931-1958, Vol. 4), and this symbol has to access a representation of the meaning of negation. Negation is therefore a counterexample to the Aristotelian thesis that everything in the mind is represented as an image. The model that is negated, however, can be iconic and rooted in a sensory modality. This possibility forms a rapprochement with those theories that base all mental representations in sensory modalities (cf. Barsalou, 1999; Glenberg et al., 1999).

Complementary predicates, such as *odd* and *even*, allow the reference of a negative assertion to be represented in a single complementary model. Sentential connectives call for complementary models in which symbolic negation also plays a role. Consider, for example, an inclusive disjunction of the grammatical form: *A or B, or both*, in which *A* and *B* are atomic clauses. This affirmative assertion refers to three possibilities, and so the fully explicit models built from the intension of the sentence are:

$$\begin{array}{l} A \quad \neg B \\ \neg A \quad B \\ A \quad B \end{array}$$

The negation of the sentence: *it is not the case that A or B or both*, refers to the complement of these three possibilities, namely:

$$\neg A \quad \neg B$$

The core meaning entails that an assertion and its negation contradict one another: one is true, and one is false, because their respective possibilities are disjoint and exhaustive. Hence, the core meaning of negation according to the present theory corresponds to its meaning in logic (see, e.g., Jeffrey, 1981). A question, such as:

11a. Is the circle on the right of the triangle?

has the true answer *no* if and only if the corresponding negative assertion, *the circle is not on the right of the triangle*, is true. Hence, the meaning of *no* in answer to such a question is to negate the affirmative proposition that is questioned, i.e., the negation has a large scope. Likewise, a negative request or command, such as:

11b. Please do not put the circle on the right of the triangle

calls for the listener to act, or to refrain from acting, so as to ensure that the corresponding negative assertion is true, i.e., the circle is not on the right of the triangle. Negative commands are prohibitions. The response *no* to such command means that the listener refuses to comply. That is, the listener does not commit to making the corresponding negative assertion true.

A sentence containing a negated noun phrase, as in:

12a. It was Viv, not Pat, who left the lights on

has the paraphrase:

12b. Viv left the lights on, and it is not the case that Pat left the lights on.

The core meaning of negation applies to the second clause in this case. However, the core interpretation of negation, or of sentential connectives, does not always occur. It can be modulated to a weaker interpretation – a matter to which we return later.

The core meaning applies to the negation of assertions containing quantifiers, such as *all artists*. There are various theories of the mental representation of determiners such as *all* and *some* (see e.g., Politzer, Van der Henst, Delle Luche, & Noveck, 2006, for a review of some diagrammatic systems, and Khemlani & Johnson-Laird, 2011, for a review of the mental representations of determiners in syllogistic reasoning). Not all of

these theories, however, generalize to assertions containing multiple quantifiers (see below) or to quantified assertions such as *more than half the women are married*, which cannot be expressed in a logic in which quantifiers range over individuals (the first-order predicate calculus). Instead, a quantifier such as *more than half the women* calls for quantification over sets. If, say, there are three women, 1, 2, and 3, in the relevant set, then the quantifier refers to one of the following sets:

{woman1 woman2}

{woman1 woman3}

{woman2 woman3}

{woman1 woman2 woman3}

and the assertion is true provided that one of these sets is included in the set of married women. The negative assertion:

13. It is not the case that more than half the women are married
calls for a member of the complement of the preceding set to be in the set of married individuals. The complement consists of these three sets:

{woman1} {woman2} {woman3}

where we ignore the empty set. If one of these sets or none of them is included in the set of married women, then the negative assertion is true. The mental model of *more than half the women are married* uses a set of tokens to represent the three women, but without specifying their identities unless they are known independently:

woman married

woman married

woman

This single mental model captures all the possibilities in the set above, because it leaves open whether or not the third woman is married. The negation of the assertion has the mental model:

woman \neg married

woman \neg married

woman

This model is inconsistent with the model of the affirmative assertion, but one of the two models must hold, because together they include all the possibilities: there are zero through three of the women who are married. Hence, the two models represent contradictory assertions. If a set has an unknown number of members, then the model represents a small but arbitrary number; and with numbers that are too large to be represented one-to-one, the model has numerical labels (Johnson-Laird, 1983, p. 443).

This treatment of quantifiers is a step towards an account of “generalized quantifiers,” but has the advantage of psychological tractability (see Barwise & Cooper, 1981; Montague, 1974; Mostowski, 1957; Partee, 1975).

The negation of quantified assertions has further complications, because the scope of negation may not be the whole clause. For example, these three different sentences can all refer to the same situation:

14a. Not all of the students know all the professors.

14b. Some of the students do not know all the professors.

14c. Some of the students do not know some of the professors.

Linguistic analyses often derive these sentences from the same underlying logical form (see, e.g., Harman, 1972; Johnson-Laird, 1970; May, 1985). According to the model

theory, however, their potential equivalence is an emergent property from models of the possibilities to which they refer. Assertion 14a is case of large scope negation. The negation in 14b has a smaller scope of just the verb phrase: *Some of the students are such that they do not know all of the professors*. The negation in 14c has a still smaller scope of just the relation: *Some of the students and some of the professors are such that the former do not know the latter*. (Readers familiar with the predicate calculus will recognize our “loglish” paraphrases of its expressions, see Jeffrey, 1981). However, a model in which some students do not know some professors is a model of all three assertions.

The core meaning extends naturally to cases in which negation has a small scope.

An assertion such as:

15a. Some guests did not arrive late

can be paraphrased as:

15b. There are some guests such that it is not the case that they arrived late.

The normal representation of negation is to use a symbol for negation with a model of the unnegated assertion within its scope, as in the model of (10) above. But, when there is a complementary predicate, a model of that can be constructed. Hence, a model of (15a) calls for a representation of guests, some of whom are in turn represented as having the negation of *arrived late*. An emergent property of this model is that it corresponds to the negation of the universal assertion, *all the guests arrived late*. The universal claim refers to three sorts of individual: guests who arrived late (necessarily in the model), non-guests who arrived late (possibly in the model), and non-guests who did not arrive late (possibly in the model). Its negation accordingly asserts the existence of the one sort of individual

who is impossible according to the original universal assertion, namely, guests who were not late. The construction of the requisite models calls for the recovery of the modal status of individuals in the unnegated assertion, i.e., those that are necessarily in the model given the truth of the assertion, and those that are necessarily *not* in the model given the truth of the assertion (cf. Stenning & Yule, 1997, p. 117). The remaining sorts of individual are possibly, or possibly not, in the model. Hence, in a model of the affirmative assertion, *some of the guests arrived late*, a necessary individual is one who is a guest and arrived late. The modal status of individuals allows the negation of a mental model to be fleshed out as a fully explicit model.

In general, the task of formulating or understanding the negation of a multiple-clause assertion should be difficult for naïve individuals. A negative assertion, such as:

16a. It is not the case both that the election is next month and that Viv has registered to vote.

should be difficult to understand, because the negation refers to the set of possibilities that are the complement of those to which the corresponding affirmative conjunction refers. The preceding assertion (16a) has the grammatical form:

16b. It is not the case that both E and R,

where *E* stands for *the election is next month*, and *R* stands for *Viv has registered to vote*.

Granted the core interpretation of negation and conjunction, (16b) refers to the following three possibilities:

$\neg E \quad \neg R$

$\neg E \quad R$

$E \quad \neg R$

A more rapid method to establish the meaning of negation would rely on prior knowledge of the negations of assertions formed with each of the main connectives, e.g., the negation of the conjunction above is, *Not E or not R, or neither*. But, naïve individuals are unlikely to know De Morgan's laws for interrelating conjunctions and disjunctions. The model theory postulates a more plausible hypothesis. Individuals think about possibilities, where a possibility consists of a conjunction of individuals, their properties, and the relations among them. Each mental model accordingly represents such a conjunction. In order to interpret the negation of a multiple-clause assertion, such as the one above, individuals envisage these models separately: they make a series of independent small scope negations of individual clauses. Hence, with *It is not the case that both E and R*, individuals begin with the possibility in which negation is applied to each clause: *not-E and not-R*. This possibility is not consistent with the original affirmative assertion, and so they realize that it is one possibility in which the negation holds. They may go no further, but if they do, they apply the negation to only one of the clauses, e.g., *not-E and R*, and appreciate that it is also not consistent with the original affirmative and accordingly a possibility in which the negation holds. And they may similarly grasp that *E and not-R* is also a possibility in which negation holds. Finally, they need to consider the case, *E and R*, which is a possibility consistent, say, with the negation of an exclusive disjunction.

The general procedure is to construct a series of models of conjunctive possibilities. It starts with negations of both clauses, and checks whether the resulting possibility is consistent with the unnegated assertion. It then negates each clause, and accepts only those possibilities that are not consistent with the unnegated assertion.

Finally, it affirms both main clauses. In each case, if a model is consistent with the unnegated assertion, it is rejected; otherwise, it is accepted as consistent with the negation. This hypothesis applies to all connectives between main clauses, but it is recursive so that it can cope with clauses within clauses. To be right for the right reasons depends on completing the full sequence of all possible conjunctions based on the two clauses.

There is an important rider to the hypothesis: individuals are likely to fail to construct the full sequence of models, which is difficult and time-consuming to envisage. Hence, they should be more likely to respond correctly if they are asked to evaluate given possibilities. In sum, naïve individuals should formulate the denial of compound assertions with multiple main clauses by envisaging, one at a time, the various sorts of possibility in which the denial holds. The order of constructing the models is unlikely to be constant, but it should usually begin with the negations of both clauses.

The possibilities that individuals envisage in interpreting negative assertions have an immediate impact on how they should formulate denials. If they are asked to deny a conjunction, *A and B*, the theory postulates that they should respond: *Not-A and not-B*, because it corresponds to the first model that they should envisage. But, the process of denying assertions with multiple clauses should also reflect the small scope heuristic. This bias should occur directly in the case of assertions with subordinate clauses. As an example, consider (17a):

17a. Before he served tea, the butler put on white gloves.

Its sentential negation has a large scope:

17b. It is not the case that before he served tea, the butler put on white gloves.

This large scope negation contradicts the original affirmative assertion, because it allows for the possibility that the butler didn't serve tea. This latter possibility can be expressed explicitly:

17c. It is not the case that before he served tea, the butler put on white gloves, because the butler didn't serve tea.

A small scope negation of (17a) is:

17d. Before he served tea, the butler did not put on white gloves.

But, does this assertion contradict the affirmative assertion (17a)? One view is that an affirmative assertion of the form *Before A, B* presupposes the truth of its subordinate clause, i.e., unless this clause is true, the sentence fails to make an assertion, and so it is neither true nor false (see, e.g., Burton-Roberts, 1989). If the negative assertion similarly presupposes the truth of its subordinate clause, then such pairs of sentences as, *Before A, B*, and, *Before A, not B*, contradict one another, granted that the presupposition, *A*, is true. But, suppose *A* is false. Some theorists argue that in that case the assertion forms a “black hole” for truth values: it swallows them up so that it has no truth value whatsoever. And this view has led in turn to accounts of natural language based on various sorts of three-valued logics, logics with truth-value gaps, and much else besides (see, e.g., Horn, 2001, Ch. 2).

The nature of presuppositions in negative assertions is highly controversial (see, e.g., Carston, 1998; Horn, 1990). Some theorists deny that they are semantic (e.g., Heinämäki, 1972; Karttunen & Peters, 1979); some theorists deny that negation is ambiguous in either meaning or logical scope (e.g., Atlas, 1977, 2004); and some theorists deny that there are truth-value gaps (e.g., Gazdar, 1979; Kempson, 1975).

Almost any combination of these views or their respective denials has its adherents. The linguistic issues may be too complicated to be addressed by psychological experiments. Yet, as we showed above, you can deny a claim such as: *Before he served tea, the butler put on white gloves* (17a), by asserting its sentential negation: *It is not the case that before he served tea, the butler put on white gloves, because the butler didn't serve tea* (17c). The final clause here blocks the implicature that the butler served tea. If your denial in (17c) is true then the assertion it denies, (17a), is false. It follows that the failure of the presupposition that the butler served tea does not create a truth-value gap in (17a). This assertion is simply false given the truth of (17c). The model theory accordingly postulates that the problem of presuppositions – at least as far as negation is concerned – can be resolved without abandoning the position that propositions have only two truth values: true or false (see, e.g., Atlas, 1977; Gazdar, 1979; Karttunen & Peters, 1979; Kempson, 1975).

The model theory has implications for the pragmatics of negation. Negations are often used to deny misconceptions (see, e.g., Jespersen, 1917; Sigwart, 1895; Wason, 1965; Brown, 1970; Givón, 1978, 1979). You would be unlikely to assert, say, that spiders are not birds, because no-one is likely to suppose that they are. But, it is plausible to assert that spiders are not insects, because many individuals are under this misconception. However, the limitations of treating negation as a speech act of denial were recognized at least as early as Frege (1919/1960). Negations have uses apart from correcting misconceptions (cf. Giora, 2006). They can be a parsimonious way to communicate information, e.g.:

18. Speaker A: How did that soccer team you coach do last season?

Speaker B: Fine; but we lost every game when the team did not have a striker. No parsimonious affirmative way exists to refer to a team without a striker. Similarly, negations can be used to introduce novel ideas. Varzi's (2008) negative biography is a testament to this fact. Here's a brief extract:

“In short: I am not what I could have been, though I could have been what I am not. I don't live in Italy but I was not born in the USA (and not in this century). I'm not dead either, thank goodness, at least not yet. And I am no longer unemployed.”

Given the difficulty of constructing models of the possibilities to which a negative assertion refers, the model theory postulates a general effect of context. It should be easier to understand a negation when individuals already have in mind the models of the corresponding affirmative assertion, i.e., it takes less time. A corollary is that explicit negations using “not” should be easier to grasp as denials than implicit negations, such as the use of a complementary predicate (e.g., “open”) to deny its antonym (e.g., “closed”). Hence, if you have already constructed a representation of the assertion that the circle is on the right of the triangle:

$$\boxed{\Delta \quad \circ}$$

then it should be straightforward to interpret an explicit denial of the assertion. You have only to introduce a symbolic negation of your model:

$$\neg \boxed{\Delta \quad \circ}$$

The process should be just as fast, if not faster, than the comprehension of the original affirmative assertion. If, instead, you are told that in fact the circle is left of the triangle, you have to infer that this assertion is an implicit negation of the previous assertion.

Kaup, Zwaan, and Lüdtke (2007) have proposed an analogous idea. They argue that negation is represented in two stages: first, the expected situation, and, second, the actual state of affairs that the negation describes. The present theory is consistent with this idea, but it diverges from it in at least one respect: it allows that negations do not always occur in a context in which there is an expected situation.

We have outlined the theory in some detail, and we now turn to its principal predictions, focusing on those that are novel, and then we evaluate the evidence in relation to each of them.

Predictions and evidence

The model theory makes five main predictions, and we organize them according to whether they concern parsing, meaning and comprehension, or context and usage. The first prediction is:

Prediction 1: The parsing of negative assertions should return a small scope for negation, where possible, because such a scope minimizes the number of mental models of assertions, and thereby reduces the processing load on working memory.

Individuals should therefore tend to formulate and to interpret negations as having small scopes. They should also find it easier – less error-prone and less time-consuming – to cope with small scope negations.

The psycholinguistic literature contains sporadic evidence that can be interpreted post hoc as corroborating prediction 1. Individuals are indeed faster to evaluate negatives that have small scopes rather than large scopes. For example, Johnson-Laird

(1970) observed that individuals are faster to classify Euler diagrams for assertions with the grammatical form *Some A are not B*, in which negation has a small scope, than for assertions with the grammatical form *Not all A are B*, in which negation has a large scope. Yet, as the participants' responses showed, they tended to interpret the two sorts of assertion as referring to the same situation. Likewise, Clark and his colleagues observed that the smaller the scope of negation the faster their participants were to understand the expression relative to the time to understand a corresponding affirmative (see, e.g., Clark, 1974). In contrast to their understanding of a denial, individuals verify an implicit negation faster than an explicit negation (Clark, 1969, 1974): negation in the predicate, "absent", has a smaller scope than in the predicate, "not present". Similarly, children appear to misremember negatively quantified assertions (e.g., *No As are Bs*) as negative generics (e.g., *As are not Bs*; see Leslie & Gelman, 2010).

Inferences from negatives also provide a test of the small-scope prediction.

Negative assertions support the following sequence of valid inferences:

not(A)

∴ not(A & B)

∴ not(A & B & C)

Each inference is valid whether *A*, *B*, *C*, are propositions or sets such as those denoted by *tall*, *dark*, and *handsome*; and the sequence is *monotone decreasing* because each successive set in the sequence has fewer members (see Barwise & Cooper, 1981). For example, if a person isn't tall, then the person isn't tall, dark, and handsome, either.

We carried out a systematic study to determine whether individuals make small-scope interpretations. The participants had to answer questions such as:

19. They're not living adult males. So, who could they be?

If negation has a large scope over the conjunction of all three predicates, there are seven possibilities: all combinations of the three predicates and their negations except for the one in which the individuals are living and adult and male. The participants were told to list as many possibilities as they could think of, but they almost always interpreted the question with a smaller scope, consistent with just over four possibilities on average.

The model theory's second prediction concerns the core meaning of negation:

Prediction 2: The core meaning of negation refers to the complement of the set to which the unnegated sentence, clause, or constituent refers.

Granted that a negation refers to a complement of a set, it follows at once that a sentential negation contradicts the corresponding affirmation. It also follows that the negation of a negation should cancel out, and refers in turn to the original set. Such cases abound in daily life. The meaning of:

20a. It's not the case that he didn't phone for the police

is that:

20b. He did phone for the police.

The accumulation of negations, however, can make comprehension difficult, e.g.:

21. It is wrong to deny that it is not the case that he didn't phone for the police.

Each negation reverses the extension of what it negates, and so comprehension calls for taking the complement of a set, and then its complement, and so on, over and over. And so it grows progressively more difficult to interpret a sequence of negations.

The grammar of certain languages and dialects can reduce multiple negations to a single negation, a phenomenon known as “negative concord” (Labov, 1972), as in the following colloquial expression:

22. I ain't got no money.

Negative concord appears to be a general phenomenon of natural language. It is found in German, Spanish, Japanese, and many other languages, and its ubiquity suggests that multiple negations are difficult to understand, and so they can serve as multiple cues to a single negation.

The core meaning of negation implies that the comprehension of negation calls for the computation of the complement of a set. The difficulty of the computation should depend on the size of the complementary set. A special case occurs with complementary predicates, such as “open”, which have only a small complementary set that has its own description, such as “closed”. They allow a negation to be translated into an affirmative, as in the translation of “not open” into “closed”. Individuals make such translations when they have to verify assertions containing complementary predicates (e.g., Clark, 1974; Wason & Johnson-Laird, 1972). Conversely, Schaeken and Schroyens and their colleagues have shown that the greater the number of elements in the complementary set, the more likely individuals are to endorse explicit negative conclusions (Schaeken & Schroyens, 2000; Schroyens, Verschueren, Schaeken, & d'Ydewalle, 2000; see also Barrouillet & Lecas, 1998). And, as Beltrán, Orenes, and Santamaría (2008) have shown, when individuals have to formulate a true sentence in place of a false sentence in a story, they are more likely to use a negative sentence when the false sentence contains a contrastive predicate, such as “green”, as opposed to a complementary predicate, such as

“even”. With the complementary predicate, they tend to assert an affirmative sentence containing its opposite, “odd”.

In a study of metaphors, Hasson and Glucksberg (2006) examined the comprehension of affirmative metaphors, such as, “my lawyer is a shark”, and their negations, “My lawyer is not a shark”. In the early stages of comprehension, both sorts of metaphor enhanced the accessibility of words related to the affirmative metaphor. After about a second, however, the affirmative assertions continued to make these words accessible, whereas the negative metaphors no longer did so. These results also corroborate the hypothesis that negative metaphors are represented at least initially by a symbolic negation of the unnegated state of affairs (see also Giora, 2006, and see Nieuwland & Kuperberg, 2008, for a comparable result using Event-Related Potentials elicited by plausible and implausible negations). In sum, the ability to cancel double negations, the phenomenon of negative concord, and the effects of the size of the complementary set, all lend credence to the core meaning of negation.

The theory’s third prediction concerns the modulation of the core meaning of negation:

Prediction 3. Meaning, reference, and knowledge can modulate the core interpretation of negation so that a negative is merely contrary to the corresponding unnegated assertion.

The prediction is a consequence of the role of modulation in the model theory, that is, the theory postulates that the meanings of clauses, their referents, and knowledge about them, can modulate the meanings of connectives in everyday language, such as *if*, *or*, and *and*, and yield interpretations outside their core logical meanings. How knowledge modulates

the interpretation of connectives has been explained in general terms: it blocks the construction of otherwise possible models of assertions (Johnson-Laird & Byrne, 2002; Orenes & Johnson-Laird, in press; Quelhas, Johnson-Laird, & Juhas, 2010). Modulation should accordingly occur with negation. If the core meaning of a negation refers to two possibilities, and modulation blocks the construction of one of them, then the negative assertion and its corresponding affirmative no longer exhaust the set of possibilities. Both assertions cannot be true, but they could both be false. In other words, the negative and affirmative assertions are only contrary to one another, not contradictory. As an example, consider the assertion:

23a. I don't have an IQ equal to 120

Its core logical interpretation is consistent with two possibilities: the speaker has an IQ of less than 120 or the speaker has an IQ of more than 120. But, the context in which the statement is made should modulate its interpretation, e.g.:

23b. Pol Pot has decreed that everyone with an IQ of at least 120 will be killed, but I don't have an IQ of 120.

A very large literature exists on this sort of phenomenon (e.g., Grice, 1989; Horn, 2004; Levinson, 2000; Sperber & Wilson, 1995), but there is a broad consensus that knowledge, including knowledge of the conventions of conversation, can modulate the core logical interpretation of negation. Evidence suggests, however, that preschool children tend to think more like logicians and are less likely to infer "scalar implicatures" of the sort illustrated in (23b) (see Noveck, 2001; Papafragou & Musolino, 2003).

The fourth prediction of the model theory concerns the effects of context on the ease of comprehending negation. Negations can be odd if they come out of the blue

without a contextual justification. Freud (1925) therefore counseled psychoanalysts to ignore negation in interpreting a patient's associations. The patient says, "You ask who this person in the dream can have been. It was *not* my mother." Freud remarks: "We amend this: so it was his mother". The patient protests too much, but it is not his claim that gives him away, but rather its lack of a plausible context. No one had supposed that the figure in his dream was his mother, until he denied it. Many theorists, as we pointed out earlier, have accordingly argued that the function of negation is to deny preconceptions. And some evidence provides corroboration for this claim. For instance, Wason (1965) proposed that it is more plausible to deny that an exception has a property of a set of stimuli than vice versa, and more plausible to deny that a smaller part of a stimulus has a property of the larger part than vice versa. His experiments called for participants to complete affirmative or negative sentences to make them true descriptions, and the results corroborated his first hypothesis, but not his second one.

The model theory makes a more general prediction about the role of context on the processing of negation:

Prediction 4: Individuals should find it easier to understand a negation if they have already constructed the models of the corresponding affirmative assertion.

A special case of this prediction occurs when the negation functions to deny a misconception, but the prediction allows that negations are not always used with such a function – as we mentioned earlier, they can be parsimonious ways to communicate information. Likewise, affirmative assertions can be used to deny negative assertions (Giora, 2006), as in:

24. Speaker A: Credit default swaps no longer exist.

Speaker B: In fact, they do still exist.

Yet, individuals should be more familiar with negative assertions that function as denials than with affirmative assertions that do so.

Studies have corroborated the prediction that negation should be understood faster when individuals have already represented the corresponding affirmative assertion (see, e.g., Glenberg et al., 1999; Johnson-Laird, 1967; Kaup et al., 2007; Lüdtke & Kaup, 2006; Nieuwland & Kuperberg, 2008; Singer, 2006; Wason, 1965; Watson, 1979).

One study has corroborated a speed-up of inferences in which individuals draw their own conclusions from a contextual denial (Johnson-Laird & Tridgell, 1972).

Premises such as:

25. Either John is intelligent or he is rich.

He is not rich.

yield the conclusion, *He is intelligent*. The inference takes longer when the second premise is an implicit negation, *He is poor*. It takes time to grasp that “poor” implies “not rich”, and this realization is an extra step in the inference from an implicit negation. But, what takes still longer is an inference from premises of the following sort, in which an affirmative premise contradicts the negative clause in the disjunction:

26. Either John is intelligent or he is not rich.

John is rich.

The participants made more errors in this case than in the others, and also took longer to draw a conclusion. Negations expressing denials may be easier to understand than affirmatives expressing denials, perhaps because denials are more often couched as negatives. But, another explanation is that the negation in the disjunction is itself difficult

to understand, because individuals have not got in mind a representation of the proposition that this negation contradicts. Either way, denials that have no preceding affirmative proposition are harder to grasp.

The fifth and final prediction of the model theory concerns the comprehension and production of the negations of compound assertions based on sentential connectives, such as *if*, *or*, and *and*. Individuals who are not logicians do not know the negations of such assertions, and so they have to enumerate them from a consideration of the possibilities to which the corresponding affirmatives refer. The following prediction is novel, and perhaps counterintuitive, and is based on the idea that the greater the number of mental models of various sorts of compound assertions, the harder it should be to understand them:

Prediction 5. Those affirmative assertions with only one mental model should be easier to understand than those with multiple mental models. Their respective negations should switch in difficulty, because the complement of one model is a set of multiple models, whereas the complement of multiple models is a set of one or two mental models.

This “see saw” effect is easy to understand in the case of compound assertions such as conjunctions and disjunctions. Two atomic propositions and their respective negations yield four possible models:

A	B
A	¬B
¬A	B
¬A	¬B

A conjunction of the form:

27. A and B

refers to only one of these models, but an inclusive disjunction of the form:

28. A or B or both

refers to the first three of them. Hence, the conjunction should be easier to understand than the disjunction. In contrast, the negation of the conjunction, *not both A and B*, refers to the three models that are the complement of the model of the original conjunction (27), whereas the negation of the disjunction, *not A or B*, refers to the one model that is the complement of the three models of the original disjunction (28). Hence, the negation of the conjunction should be *harder* to understand than the negation of the disjunction. This predicted interaction hinges, of course, on the theory that individuals construct mental models of assertions, and on the core meaning of negation. Theories in which models of possibilities play no part are unlikely to make the prediction (cf., e.g., Braine & O'Brien, 1998; Rips, 1994).

Granted that individuals do not know the possibilities to which the negations of compound assertions refer, they have to enumerate their possibilities. We illustrated this procedure in our earlier account of the theory. For the negation of a conjunction, *It is not the case that both A and B*, it yields the following possibilities, though individuals may not proceed beyond the first possibility:

$\neg A \quad \neg B$

$\neg A \quad B$

$A \quad \neg B$

Those who are presented with these cases should grasp that each is contrary to the

conjunction. The final step of the procedure concerns the model:

$$A \quad B$$

It is obviously consistent with the unnegated assertion, and so it is rejected. The negation of an inclusive disjunction, *It is not the case that A or B or both*, yields as a first model:

$$\neg A \quad \neg B$$

which is inconsistent with the unnegated assertion, and so it holds for the negated assertion. Indeed, it represents the only sort of possibility in which the negation holds. Hence, the theory predicts that the negation of an inclusive disjunction, which yields only one model, should be easier than the negation of a conjunction, which yields three models. The prediction was corroborated by two recent studies carried out in our laboratory. When participants had to state what was possible given denials of conjunctions and disjunctions, the negated conjunctions yielded 18% correct responses whereas the negated disjunctions yielded 89% correct responses. Likewise, when they had to formulate denials of conjunctions and disjunctions, they made correct denials for 0% of conjunctions but for 67% of inclusive disjunctions (Khemlani, Orenes, & Johnson-Laird, in preparation).

Consider the denial of a conditional assertion, such as:

29a. If Obama wins Ohio in 2012 then he will win the Presidential election

Individuals who deny such an assertion are likely to do so using a small scope for the negation:

29b. If Obama wins Ohio in 2012 then he won't win the Presidential election.

The meaning of conditionals is highly controversial (see, e.g., Evans, 2007; Handley et

al., 2006; Johnson-Laird et al., 2009), but studies of conditionals have shown that many individuals make the small scope interpretation of conditional assertions (Handley, Evans, & Thompson, 2006). According to the model theory, however, some individuals should make a large scope denial. The small-scope formulation is, in fact, not logically correct, because a pair of assertions of the form, *If A then B*, and, *If A then not B*, do not contradict one another. They are both compatible with the possibility that A is false. The error is transparent in the case of an assertion, such as (30a):

30a. If they're Democrats then they are honest.

Its small scope negation is (30b):

30b. If they're Democrats then they are not honest.

These two assertions are contraries, i.e., they can't both be true, but they can both be false. Indeed, they are both false, granted that some of the Democrats are honest, and some are not. The logical negation of *if A then B* is accordingly, *A and not-B*. This conjunction contradicts the conditional. And it fits the robust findings of Barrouillet and his colleagues (e.g., Barrouillet, Grosset, and Lecas, 2000) that adults list three cases as possible given a basic conditional, *If A then B*:

A	B
¬A	B
¬A	¬B

Hence, only one case is impossible:

A ¬B.

The model theory accordingly allows that some individuals should realize that the negation of a conditional is: *A and not-B*, because they recognize that this case falsifies

the conditional (Johnson-Laird & Tagart, 1969; Oaksford & Stenning, 1992). Studies of conditionals have shown that many individuals make the small scope interpretation of the negation, but some individuals make the correct large scope interpretation (see Johnson-Laird, Byrne, & Girotto, 2009; Girotto & Johnson-Laird, 2004).

General Discussion

Negation is so commonplace in everyday language that few people are aware of its problematic nature. In this article, we have presented a comprehensive model-based theory of the main aspects of understanding and formulating negative assertions. The theory has nothing to say about many philosophical problems, such as whether there are negative facts. Its focus is instead on those aspects of negation that are open to empirical investigation.

The theory postulates that individuals tend to interpret, and to formulate, negations and other similar operators as having a small scope. Some prior evidence bore out this hypothesis, and we described a recent unpublished experiment that corroborated it in a systematic study of the interpretation of assertions such as, *They're not living adult males*. The core meaning of negation is a function that takes a single argument, which refers to a set of models of possibilities, and returns the complement of the set. The core meaning accordingly abides by the Aristotelian constraint that the negation of an assertion contradicts the corresponding affirmative assertion (see *De Interpretatione* in Aristotle, 1984, Vol. 1). This constraint explains the cancellation of double negations, i.e., an assertion, such as, *He is not an unbeliever*, is synonymous with *He is a believer*.

The interpretation of negation can be modulated by meaning and context when knowledge blocks models of possibilities (see Johnson-Laird & Byrne, 2002; Quelhas et al., 2010). As a result, modulation yields only a contrary instead of a contradiction to the corresponding affirmative assertion. Linguists have long known about this phenomenon (see Horn, 2001, p. 370), which is illustrated in the following “scalar implicature”:

31. Speaker A: How long have you been here?

Speaker B: Not for ten minutes.

A logical interpretation of Speaker B’s remark is consistent with two possibilities: he has been here for less than ten minutes or for more than ten minutes. But, the context of the dialog is likely to block one of these possibilities, and to leave only the one in which the speaker has been here for less than ten minutes.

A common use of negation is to correct misconceptions, and some authors have argued that such denials are its sole use (e.g., Wason, 1965). In fact, negation is sometimes the only parsimonious way in which to express a proposition – a proposition that in no way corrects anyone’s misconceptions, e.g.:

32. Speaker A: Have you seen any of Hitchcock’s movies?

Speaker B: I haven’t seen *Marnie*, but I’ve seen all the rest.

The theory accordingly makes the more general prediction that a negation should be easier to understand when individuals have already represented the corresponding affirmative assertion, and abundant experimental evidence corroborates this prediction (e.g., Kaup et al., 2007).

Finally, the theory postulates that individuals usually do not know the negations corresponding to the different sentential connectives, but instead have to construct them

on the fly. They consider a sequence of conjunctive models of possibilities, checking them against the corresponding affirmative assertion. It follows that an inclusive disjunction, *A or B or both*, should be easy to deny, because the first model that individuals should consider is the one true negation of the disjunction: *not-A and not-B*. In contrast, a conjunction, *A and B*, should be difficult to deny, because its denial is equivalent to *not-A or not-B or neither*, and so individuals need to envisage fully explicit models of three sorts of possibility. Conditionals with the structure *if A then B* are an intermediate case. Their denials should contradict them, and *A and not-B* should be more difficult to envisage because this model is unlikely to be the first one that comes to mind. Their denials should also be susceptible to a small scope interpretation, and so some individuals should deny a conditional by using another conditional: *if A then not-B*. But, some individuals do take *A and not-B* to be the denial of a conditional, and likewise most people take this denial as sufficient to falsify a conditional too (Evans, Newstead, & Byrne, 1993; Johnson-Laird & Tagart, 1969; Oaksford & Stenning, 1992).

We could describe the world without the use of negation: we could use “false” instead. Yet, to do so in a natural language leads to unnecessary verbosity; negation is a convenient way to deny propositions and to describe the world parsimoniously. The heart of the model theory is that the meaning of a negation refers to a set of models complementary to those of the corresponding affirmative. The smaller the scope of a negation, the smaller the number of these models, and so individuals are biased to reduce the scope. This hypothesis is one of the major novel components of the theory that we have described here. Another such component is the great difficulty that we all have in comprehending the negations of compound sentences containing sentential connectives.

Easy affirmations yield difficult negations.

References

- Aristotle (1984). *The complete works of Aristotle*, Vols. 1 and 2. J. Barnes (Ed).
Princeton, NJ: Princeton University Press.
- Atlas, J.D. (1977). Negation, ambiguity, and presupposition. *Linguistics and Philosophy*,
1, 321-336.
- Atlas, J.D. (2004). Presupposition. In L.D. Horn & G. Ward (Eds.), *Handbook of
pragmatics* (pp. 29-52). Oxford: Blackwell.
- Barrouillet, P., Grosset, N., & Lecas, J.F. (2000). Conditional reasoning by mental
models: chronometric and developmental evidence. *Cognition*, 75, 237-266.
- Barrouillet, P., & Lecas, J.F. (1998). How can mental models theory account for content
effects in conditional reasoning? A developmental perspective. *Cognition*, 67,
209-253.
- Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22,
577-660.
- Barwise, J. (1989). *The situation in logic*. Stanford, CA: Center for the Study of
language and Information.
- Barwise, J. (1993). Everyday reasoning and logical inference. *Behavioral and Brain
Sciences*, 16, 337-338.
- Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language.
Linguistics and Philosophy, 4, 159-219.
- Beltrán, D., Orenes, I., & Santamaría, C. (2008). Context effects on the spontaneous

- production of negation. *Intercultural Pragmatics*, 5-4, 409-419.
- Birney, D., & Halford, G. S. (2002). Cognitive complexity of suppositional reasoning: An application of relational complexity to the knight-knave task. *Thinking & Reasoning*, 8, 109-134.
- Bock, J.K. & Mazella, J.R. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition*, 11, 64-76.
- Braine, M.D.S., & O'Brien, D.P., Eds. (1998). *Mental logic*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Brown, R. (1970). Personality and style in Concord. Ch. 13, *Psycholinguistics: Selected papers* (pp. 336-392). New York: The Free Press.
- Burton-Roberts, N. (1989). On Horn's dilemma: Presupposition and negation, *Journal of Linguistics*, 25, 95-125.
- Carlson, K., Frazier, L., & Clifton, C. (2009). How prosody constrains comprehension: A limited effect of prosodic packaging. *Lingua*, 119, 1066-1082.
- Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, 82, 45-73.
- Carston, R. (1998). Negation, presupposition and the semantics/pragmatics distinction. *Journal of Linguistics*, 34, 309-350.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Clark, H.H. (1969). Linguistic processes in deductive reasoning. *Psychological Review*, 76, 387-404.
- Clark, H.H. (1974). Semantics and comprehension. In T.A. Sebeok (Ed.) *Current trends*

- in linguistics, Vol. 12: Linguistics and adjacent arts* (pp. 1291-1428). The Hague: Mouton.
- Clark, H.H., & Chase, W.G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 3, 472-517.
- Dahl, O. (1979). Typology of sentence negation. *Linguistics*, 17, 79-106.
- Dale, R. & Duran, N. D. (2011). The cognitive dynamics of negated sentence verification. *Cognitive Science*, 35, 983-996.
- Evans, J.St.B.T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. Hove and New York: Psychology Press.
- Evans, J.St.B.T., Newstead, S.E., & Byrne, R.M.J. (1993). *Human reasoning: The psychology of deduction*. Hillsdale, NJ: Erlbaum.
- Frege, G. (1919/1960). Negation. In P. Geach & M. Black, *Translations from the Philosophical Writings of Gottlob Frege*. Oxford, UK:
- Freud, S. (1925). Negation. In J. Strachey (Ed.), *Complete psychological works of Sigmund Freud, Vol. 19: The ego and the id and other works* (pp. 181-185). London: Hogarth.
- Gazdar, G. (1979). *Pragmatics: Implicature, presupposition and logical form*. New York: Academic Press.
- Giroto, V., & Johnson-Laird, P.N. (2004). The probability of conditionals. *Psychologia*, 47, 207-225.
- Givón, T. (1978). Negation in language: Pragmatics, function, ontology. *Syntax and Semantics*, 9, 69-105.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

- Glenberg, A. M., Robertson, D. A., Jansen, J. L., & Glenberg, M.C.J. (1999). Not propositions. *Journal of Cognitive Systems Research*, 1, 19-33.
- Giora, R. (2006). Anything negatives can do affirmatives can do just as well, except for some metaphors. *Journal of Pragmatics*, 38, 981-1014.
- Goodwin G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review*, 112, 468-493.
- Goodwin, G., & Johnson-Laird, P.N. (2006). Reasoning about the relations between relations. *Quarterly Journal of Experimental Psychology*, 59, 1047-1069.
- Goodwin, G.P., & Johnson-Laird, P.N. (2008). Transitive and pseudo-transitive inferences. *Cognition*, 108, 320-352.
- Gough, P.B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning and Verbal Behavior*, 4, 107-111.
- Grice, H.P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Haegeman, L. (1995). *The Syntax of Negation*. Cambridge University Press, Cambridge.
- Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 21, 803-831.
- Handley, S.J., Evans, J. St. B.T., & Thompson, V.A. (2006). The negated conditional: A litmus test for the suppositional conditional? *Journal of Experimental Psychology: Learning, Memory, Cognition*, 32, 559-569.
- Harman, G. (1972). Deep structure as logical form. In D. Davidson & G. Harman (Eds.), *Semantics of natural language* (pp. 25-47). Dordrecht, Holland: Reidel.

- Hasson, U., & Glucksberg, S. (2006). Does understanding negation entail affirmation? An examination of negated metaphors. *Journal of Pragmatics*, 38, 1015–1032.
- Heinämäki, O. (1972). Before. In P. M. Peranteau et al. (Eds.), *Papers from the 8th Regional Meeting, Chicago Linguistic Society* (pp. 139-151). Chicago, IL: Chicago Linguistic Society.
- Horn, L. (1990). Showdown at truth-value gap: Burton-Roberts on presupposition. *Journal of Linguistics*, 26, 483-503.
- Horn, L. R. (2001). *A natural history of negation*. Rev. ed. Stanford, CA: CSLI Publications.
- Horn, L.R. (2004). Implicature. In L.R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 3-28). Oxford: Blackwell.
- Hunt, E., & MacLeod, C.M. (1978). The sentence-verification paradigm: A case study of two conflicting approaches to individual differences. *Intelligence*, 2, 129-144.
- Jeffrey, R.C. (1981). *Formal logic: Its scope and limits*. Second edition. New York: McGraw-Hill.
- Jespersen, O. (1917). *Negation in English and other languages*. Copenhagen: A.F. Host.
- Johnson-Laird, P.N. (1967). One pragmatic factor governing the use of the English language. University of London: Unpublished doctoral dissertation.
- Johnson-Laird, P.N. (1970). The interpretation of quantified sentences. In W.J.M. Levelt & G.B. Flores d'Arcais (Eds.), *Advances in psycholinguistics* (pp. 347-372). Amsterdam: North Holland.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

- Johnson-Laird, P. N. (2002). Peirce, logic diagrams, and the elementary operations of reasoning. *Thinking and Reasoning*, 8, 69-85.
- Johnson-Laird, P. N. (2006). *How we reason*. New York: Oxford University Press.
- Johnson-Laird, P.N., & Byrne, R.M.J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109, 646-678.
- Johnson-Laird, P.N., Byrne, R.M.J., & Girotto, V. (2009). The mental model theory of conditionals: A reply to Guy Politzer. *Topoi*, 28 (1), 75-80.
- Johnson-Laird, P. N., & Tagart, J. (1969). How implication is understood. *American Journal of Psychology*, 82, 367-373.
- Johnson-Laird, P. N., & Tridgell, J. (1972). When negation is easier than affirmation. *Quarterly Journal of Experimental Psychology*, 24, 87-91.
- Karttunen, L., & Peters, S. (1979). Conventional implicature. In C.-K. Oh & D. Dineen (Eds.), *Syntax and semantics, 11: Presupposition* (pp. 1-56). New York: Academic Press.
- Kaup, B., Zwaan, R. A., & Lüdtke, J. (2007). The experiential view of language comprehension: How is negation represented? In F. Schmalhofer & C. A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 255-288). Mahwah, NJ: Lawrence Erlbaum.
- Kempson, R. (1975). *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.
- Khemlani, S., & Johnson-Laird, P.N. (2009). Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, 37 (5), 615-623.

- Khemlani, S., & Johnson-Laird, P.N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, in press.
- Khemlani, S., Lotstein, M., & Johnson-Laird, P.N. (under review). A unified theory of syllogistic reasoning.
- Khemlani, S., Orenes, I., & Johnson-Laird, P.N. (in preparation). The negation of conjunctions and disjunctions. Manuscript in preparation.
- Klima, E.S. (1964). Negation in English. In J.A. Fodor & J.J. Katz (Eds.), *The structure of Language* (pp. 246-323). Englewood Cliffs, NJ: Prentice Hall.
- Knauff, M., & Johnson-Laird, P.N. (2002). Imagery can impede inference. *Memory & Cognition*, 30, 363-371.
- Labov, W. (1972). Negative attraction and negative concord in English grammar. *Language*, 48, 773-818.
- Leslie, S.-J. & Gelman, S. (2010). Quantified statements are recalled as generics: Evidence from preschool children and adults. Manuscript submitted for publication.
- Levinson, S.C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Lüdtke, J. & Kaup, B. (2006). Context effects when reading negative and affirmative sentences. In R. Sun (Ed.) *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1735). Austin, TX: Cognitive Science Society.
- May, R. (1985). *Logical form: Its structure and derivation*. Cambridge, MA: MIT Press.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press. Cambridge: Cambridge University Press.

- Montague, R. (1974). *Formal philosophy: Selected papers*. New Haven: Yale University Press.
- Mostowski, A. (1957). On a generalization of quantifiers. *Fundamenta Mathematicae*, 44, 12-36.
- Moxey, L.M., Sanford, A.J., & Dawydiak, E.J. (2001). Denials as controllers of negative quantifier focus. *Journal of Memory and Language*, 44, 427-442.
- Nieuwland, M.S., & Kuperberg, G.R. (2008). When the truth isn't too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science*, 19, 1213-1218.
- Noveck, I.A. (2001). When children are more logical than adults: Investigations of scalar implicature. *Cognition*, 78, 165-188.
- Oaksford, M., & Stenning, K. (1992). Reasoning with conditionals containing negated constituents. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 835-854.
- Orenes, I., & Johnson-Laird, P.N. (in press). Logic, models, and paradoxical inferences. *Mind and Language*.
- Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition*, 86, 253-282.
- Partee, B. H. (1975). Montague grammar and transformational grammar. *Linguistic Inquiry*, 6, 203-300.
- Peirce, C. S. (1931-1958). *Collected papers of Charles Sanders Peirce*. (8 Vols.; C. Hartshorne, P. Weiss, & A. Burks, Eds.). Cambridge, MA: Harvard University Press.

- Politzer, G., Van der Henst, J-B., Delle Luche, C., & Noveck, I.A. (2006). The interpretation of classically quantified sentences: A set-theoretic approach. *Cognitive Science*, 4, 691-723.
- Quelhas, A.C., Johnson-Laird, P.N., & Juhas, C. (2010). The modulation of conditional assertions and its effects on reasoning. *Quarterly Journal of Experimental Psychology*, iFirst, 1-24.
- Rips, L.J. (1994). *The psychology of proof*. Cambridge, MA: MIT Press.
- Russell, B.A.W. (1905). On denoting. *Mind*, 14, 479-493.
- Schaeken, W., & Schroyens, W. (2000). The effect of explicit negations and of different contrast classes on conditional syllogisms. *British Journal of Psychology*, 91, 533-550.
- Schroyens, W., Verschueren, N., Schaeken, W. & d'Ydewalle, G. (2000). Conditional reasoning with negations: Implicit versus explicit affirmation or denial and the role of contrast classes. *Thinking & Reasoning*, 6, 221-251.
- Sigwart, C. (1895). *Logic, Vol. I. 2nd*. Ed. New York: Macmillan.
- Singer, M. (2006). Verification on text ideas during reading. *Journal of Memory and Language*, 54, 574-591.
- Slobin, D.I. (1966). Grammatical transformations in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior*, 5, 219-227.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell. (Original edition published in 1986.)
- Stenning, K. & Van Lambalgen, M. (2008). *Human reasoning and cognitive science*. MIT Press.

- Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive Psychology*, 2, 239-289.
- Varzi, A. (2008). Negative biography. Achille C. Varzi Home Page. Retrieved August 24, 2010, from <http://www.columbia.edu/~av72/bio.html>.
- Wason, P.C. (1959). The processing of positive and negative information. *Quarterly Journal of Experiment Psychology*, 11, 92-107.
- Wason, P.C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology*, 52, 133-142.
- Wason, P.C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, 4, 7-11.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Wason, P.C., & Jones, S. (1963). Negatives: denotation and connotation. *British Journal of Psychology*, 54, 299-307.
- Watson, J.M. (1979). Referential description by children in negative form. *British Journal of Psychology*, 70, 199-204.
- Wittgenstein, L. (1953). *Philosophical investigations*. Trans. by G.E.M. Anscombe. New York: Macmillan.