Available online at www.sciencedirect.com

SCIENCE DIRECT®

Cognitive Psychology 50 (2005) 159–193

Cognitive Psychology

www.elsevier.com/locate/cogpsych

Naïve deontics: A theory of meaning, representation, and reasoning

Monica Bucciarelli[a,1], P.N. Johnson-Laird[b,*]

[a] *Centro di Scienza Cognitiva, Università di Torino, Torino 10123, Italy*
[b] *Department of Psychology, Princeton University, Princeton, NJ 08544, USA*

Accepted 4 August 2004
Available online 8 October 2004

Abstract

Deontic assertions concern what one ought to do, may do, and ought not to do. This paper proposes a theory of their meanings and of how these meanings are represented in mental models. The meanings of deontic assertions refer to sets of permissible and impermissible states. An experiment corroborated the ability of individuals to list these states. The most salient were those corresponding to the mental models of the assertions. When individuals reason, they rely on mental models, which do not make all states explicit. The theory predicts the most frequent conclusions drawn from deontic premises. It also predicts the occurrence of illusory inferences from assertions of permission, i.e., inferences that seem highly plausible but that are in fact invalid. Assertions of prohibitions, according to the theory, should reduce the illusions. Further experiments corroborated these predictions.

Keywords: Deductive reasoning; Deontics; Mental models; Semantics; Illusory inferences

1. Introduction

Deontic principles concern what is obligatory, permissible, and impermissible. They lie at the heart of human social relations, and underlie all ethical, legal, and

* Corresponding author. Fax: +1 609 258 1113.
E-mail addresses: monica@psych.unito.it (M. Bucciarelli), phil@princeton.edu (P.N. Johnson-Laird).
[1] Fax: +011 8159039.

0010-0285/$ - see front matter
doi:10.1016/j.cogpsych.2004.08.001

religious systems. Cultural institutions from politics to games are likewise governed by deontic rules, implicit or explicit. Hence, deontic assertions include everything from strong moral principles, such as *Thou shalt not kill*, to casual recommendations, such as *If you're in Venice*, *then you ought to see the Accademia*. Likewise, as these two examples illustrate, some deontic assertions are categorical, whereas others are conditional. Thus, deontic reasoning occurs in a variety of situations in daily life, ranging from deciding a course of action (e.g., what to do in a situation where your mother wants you to visit her, but your boss wants you to go to Denver), deciding whether an action is permissible in a complicated institutional case (e.g., whether the USA has the right to wage war without UN approval), and evaluating whether someone else's action is permissible (e.g., giving a heart transplant to a convicted murderer at the taxpayer's expense). Deontic principles vary in the rigor that they are enforced, and in their consequences, both for abiding by them and for violating them (Evans & Twyman-Musgrove, 1998). Their formulation is often a result of reasoning about matters of morality and systems of values (see, e.g., Kohlberg, 1986; Walker, 1988). This paper is not about the specific contents of deontic rules, not about who is entitled to utter deontic prescriptions to whom, and not about the motives and reasoning that leads to deontic propositions (Green, McClelland, Muckli, & Simmons, 1999; Ray, Reynolds, & Carranza, 1989; Staller & Petta, 2001). What it seeks to establish is the rudimentary foundations of any deontic system. It aims to answer three fundamental questions.

The first question is: what do deontic assertions mean? That is, what is their meaning to naïve individuals, i.e., to individuals who are not experts on deontics? For instance, a typical deontic assertion is:

(1) The boss permitted you to leave early.

The subject of the sentence—the boss—must have acted in some way to grant permission, most likely by carrying out an appropriate speech act (Searle, 1969). One condition for a felicitous act of granting permission is that the speaker has the power to grant it and to withhold it. Otherwise, no real permission exists. We subsume example 1 under a general analysis that interrelates two propositions in the following way:

(2) The boss's approval permitted you to leave early.

The scope of the paper is accordingly any assertion that has an interpretation in common with one of four sorts of deontic relation:

(3) A permits B.
A obligates B.
A prohibits B.
A permits not B.

where A stands for a proposition, such as: the boss approved the request, which refers to a possible state, and B stands for a proposition, such as: you left early, which refers to a permissible or impermissible state. The four preceding relations are basic, because any deontic discourse can be constructed from them; and they exhaust the set of possible dyadic deontic relations. Many assertions of different surface forms have the same deontic interpretation, e.g.:

(4) If you have the boss's approval then you may leave early.
You are allowed to leave early because you have the boss's approval.
You have the boss's approval and so you can leave early.

But, at the heart of any set of deontic concepts are permission, obligation, and prohibition. With them, any deontic principles can be formulated; without them, deontic discourse is impossible.

Our study focuses on the core meanings of deontic verbs. These meanings are independent of the domain under discussion. Thus, the verb 'to obligate' has a core meaning that is independent of the sentences in which it occurs, e.g.:

(5) If you borrow money then you are obligated to repay it.
Your signature on the contract implies that you are obligated to pay the fee.
If your serve touches the net cord, then you are obligated to serve again.

Each of these sentences can also be paraphrased using the modal auxiliary verb, 'must' in place of 'are obligated to.' The other verbs expressing dyadic relations are similar in that they too have core meanings. Of course, this claim is a hypothesis, which is part of the theory that we will advance and test. This theory of the meaning of deontic assertions takes as primitive the concept of a permissible state. But, deontic assertions, such as *A permits B*, are analogous to causal assertions, such as *A enables B to occur*. The deontic assertion conveys what is permissible, whereas the causal assertion conveys what is possible (see also Klaczynski & Narasimhani, 1998; and Goldvarg & Johnson-Laird, 2001). In fact, as we will see, both possibility and permissibility are implicated in the analysis of deontic meanings.

The second question that the paper addresses is: how do individuals mentally represent the meaning of deontic assertions? For example, the assertion:

(6) Promising obligates you to pay.

refers to a situation in common with one referred to by the assertion:

(7) Not promising permits you not to pay.

But, according to the theory that we present, the two assertions differ in their mental representations. The theory is based on mental models, and it postulates that assertion 6 has a model that makes salient the case in which you promise and are obligated to pay, but this case is less salient in the models of assertion 7.

The third question that the paper addresses is: how do people reason from deontic assertions? Psychological studies of this topic have focused on a particular task, the so-called 'selection task' devised by Wason (1966). There have been few studies of the conclusions that individuals draw spontaneously from deontic assertions (but cf. Quelhas & Byrne, 2003), and so the paper reports studies of such basic inferences as:

(8) Earning a salary obligates you to pay your Social Security.

Paying your Social Security prohibits you from receiving a pension from abroad.
Therefore, earning a salary prohibits you from receiving a pension from abroad.
This inference has the form:

(9) A obligates B.
B prohibits C.
Therefore, A prohibits C.

where A is the proposition that you earn a salary, B is the proposition that you pay your Social Security, and C is the proposition that you receive a pension from abroad. The inference is *valid*, i.e., if its premises are true then, as people assent, its conclusion must be true too. But, what are the mental processes that enable you to understand the deontic terms that occur in this inference, to evaluate it, and to construct it in the first place? Logicians have formulated deontic logics that capture the implications among deontic propositions. But, to the best of our knowledge, no previous psychological theory explains how individuals draw the preceding inference.

Readers might suppose that the present paper takes a narrow view of deontics. They might argue that a theory of deontic reasoning should explain how individuals deal with competing obligations, how they pick out violators of a deontic rule, how they decide whether permission for some action has been granted, why some deontic constraints are stronger than others, what determines who can make felicitious deontic assertions, and why individuals abide by some obligations more often than they abide by others. All these questions, however, presuppose an account of what deontic assertions mean, how they are mentally represented, and how inferences are drawn from them. For this reason, the present paper proposes a theory of the foundations of deontics. Answers to these other more sophisticated questions must depend ultimately on such a theory. If our theory is correct, then all deontic relations are founded on the concept of permissibility. This foundation may be narrow, but it is necessary.

In what follows, the paper sets the scene with a survey of deontic logic, the psychological literature on deontics, and the theory of mental models (Part 2). The plan of the paper then mirrors its principal goals. It proposes a theory of the core meanings of deontic assertions and of how they are mentally represented (Part 3). It reports a test of this account (Part 4). It describes how mental models underlie deontic reasoning (Part 5). It shows how their use leads reasoners into systematic fallacies (Part 6). Finally, it draws some general conclusions about deontic reasoning (Part 7).

2. Deontic logics, psychology, and mental models

2.1. Deontic logics and reasoning

Logicians have formulated various deontic logics (e.g., Hilpinen, 1971; Lewis, 2000; van Fraasen, 1973). These logics are typically based on standard logic, but introduce additional axioms that concern what is *permissible* and what is *obligatory*, which they treat as akin to the 'modal' notions of possibility and necessity. The different deontic logics have in common the following central axiom, which we paraphrase in everyday English:

(10) Whatever is obligatory is permissible.

The logics also postulate two principal dyadic operators:

(11) Obligatory (A | B): A is obligatory given B.
Permissible (A | B):A is permissible given B.

where A and B refer to propositions. These two operators are interdefinable, as are the corresponding operators of *necessity* and *possibility*, and we can paraphrase the definitions in everyday English

(12) A is obligatory given B $=_{\text{def}}$It is not permissible for not-A given B.
A is permissible given B $=_{\text{def}}$It is not obligatory for not-A given B.

Prohibition in turn can be defined as follows:

(13) A is prohibited given B $=_{\text{def}}$It is obligatory for not-A given B.

It is easy to suppose that these equivalences break down in daily life. Consider this assertion, based on an example suggested by a reviewer:

(14) The hospital's rules prohibit you from having visitors.

What are you obligated not to do? The answer is: you are *not* to have visitors. As this example illustrates, there are ambiguities in the relevant quantifier here: are you prohibited from having just one visitor, many visitors, or any visitors? Such ambiguities, however, have nothing to do with deontics. They occur in any assertion containing simple plural nouns, e.g., 'The hospital welcomes visitors.'

Formal logics enable you to construct proofs, but they do not specify the meanings of expressions. However, logicians have formulated so-called 'possible worlds' semantics, which state the conditions in which assertions are true in a given world, such as the actual world, in terms of their truth or falsity in other possible worlds. Hence, an assertion of the form:

(15) A is possible

is taken to be true in the actual world provided that A is true in at least one possible world envisageable from the actual world. This approach, which goes back to Leibniz, may seem to be circular because what is possible is treated as true in a possible world. But, there is no real circularity: an operator in a formal language—the sentential operator of possibility—is given a semantic analysis in terms of an abstract mathematical model based on sets of 'possible worlds' (Hughes & Cresswell, 1984). Within this framework, Lewis (2000) proposed analyses of *permissible* and *obligatory* in terms of possible worlds that are perfect from the standpoint of deontics. If you are obligated to do X, then in such a perfect world you do X.

Deontic logics are not intended to offer plausible psychological accounts of naïve reasoning. Nevertheless, they have several important lessons for psychology. The deontic concepts of obligation and permissibility are closely related to the concepts of necessity and possibility, but, as logic shows, a crucial difference exists between the deontic and factual domains. What is necessarily the case *is* the case; but what is obligatorily the case may not be the case: individuals have been known to fail to carry out their obligations. The two deontic operators are interchangeable even in daily life. If, for example, you are not permitted to smoke then you are obligated not to smoke. Many deontic relations are dyadic. Hence, an action can create an obligation to do something, or it can make another action permissible.

2.2. Psychological studies of deontics

Psychologists and linguists have investigated the meanings of deontic terms such as *may* and *must* (see, e.g., Steedman, 1977). They have argued that such modal auxiliaries have both an epistemic interpretation as in, 'It may rain,' and a deontic interpretation as in, 'You may smoke.' Children acquire the deontic interpretation earlier than the epistemic one, presumably because what they are allowed to do is more salient to them than what is possible (see, e.g., Lyons, 1977; Shepard, 1982). They begin to make spontaneous deontic assertions before they are 3 years old, and as they develop they gradually master the distinctions amongst deontic verbs. For example, Day (1996; see also Day & Caron, 1991) examined 6- to 13-year-old children, and showed that as they grow up they gradually differentiate the deontic meanings of the French verbs *pouvoir* (*can*) and *devoir* (*must*). Likewise, an observation that informs our analysis of prohibitions is that children seem to learn what they should *not* do earlier than they learn what they should do (see Gralinski & Kopp, 1993). Children do not grasp the real distinction between what is permissible and what is obligatory until the age of five or six (Byrnes & Duff, 1989). By then, their use of modal verbs reveals that they can conceive of actions on the part of morally responsible agents and reflect on which of them are permissible and which of them are obligatory. By this age, children are also supposed to share adult intuitions about what is physically possible and physically impossible (see Kalish, 1998; Noveck, Ho, & Sera, 1996). Hirst & Weil (1982) have suggested that the deontic interpretation of modal auxiliaries may derive from an underlying epistemic meaning, or that both interpretations may derive from a common underlying interpretation (see also Johnson-Laird, 1978).

Cheng & Holyoak (1985) proposed a major theory of deontics. Its primary aim was to explain the facilitatory effects of deontic content on Wason's (1966) selection task. The theory postulates that knowledge is represented in the form of *pragmatic reasoning schemas*. These schemas consist of sets of conditional rules. But, conditional rules in turn have universal computational power, and so no empirical observations are likely to refute schemas in general. In what follows, we therefore focus on Cheng and Holyoak's specific proposals that are empirically testable. They argue that each schema consists of a set of general rules defined in terms of classes of actions and preconditions. Here are their two main schemas:

(16) Schema for permission:
Rule 1: If the action is to be taken, then the precondition must be satisfied.
Rule 2: If the action is not to be taken, then the precondition need not be satisfied.
Rule 3: If the precondition is satisfied, then the action may be taken.
Rule 4: If the precondition is not satisfied, then the action must not be taken.

(17) Schema for obligation:
Rule 1. If the precondition is satisfied, then the action must be taken.
Rule 2. If the precondition is not satisfied, then the action need not be taken.
Rule 3. If the action is to be taken, then the precondition may have been satisfied.
Rule 4. If the action is not to be taken, then the precondition must not have been satisfied.

A deontic conditional, such as, 'If a person is drinking beer then that person must be over 19 years of age,' maps onto rule 1 of the permission schema in the following way:

(18) If the action (of drinking beer) is to be taken, then the precondition (of being over 19 years of age) must be satisfied.

Hence, the conditional assertion evokes the permission schema, and its entire set of rules becomes available to guide the process of reasoning.

Pragmatic schemas account for performance in the deontic selection task (Cheng & Holyoak, 1985). One problem, however, is that the rules in schemas are stated using the modal auxiliaries 'may' and 'must.' A semantics for deontic terms needs to explain the meanings of these terms, which are treated as unanalyzed primitives in the schemas. Moreover, these modal auxiliaries are systematically ambiguous between an epistemic interpretation and a deontic interpretation (see, e.g., Johnson-Laird & Byrne, 1996). Individuals may use these pragmatic schemas in reasoning, but they cannot be the foundation of deontics, because they do not state the deontic meanings of 'may' and 'must' but presuppose them.

Evolutionary psychologists, such as Cosmides (1989) & Cummins (1996), have postulated that individuals rely on innate modules for deontic reasoning. But, skeptics have argued that there is no need to invoke such inferential modules (e.g., Holyoak & Cheng, 1995). There are also phenomena that the proposed modules cannot explain (see Cheng & Holyoak, 1989; Johnson-Laird & Byrne, 2002; Manktelow & Over, 1990, 1991, 1995). Unfortunately, evolutionary theorists do not describe how the modules work, what obligation and permission mean, or how individuals draw conclusions from deontic assertions.

Deontic materials certainly appear to improve performance in the selection task. The burden of logical analyses (see the previous section) is that the meaning of deontic and factual assertions have a crucial difference in meaning. According to Stenning & van Lambalgen (2004), this contrast affects how individuals interpret the selection task and yields their better performance with deontic content. It can improve other sorts of reasoning (Schaeken & d'Ydewalle, 1996), but sometimes has no reliable effects (Bucciarelli & Johnson-Laird, 1999; Yang & Johnson-Laird, 2000a).

The theories that we have described so far treat deontic reasoning as dependent on special mechanisms (cf. Almor & Sloman, 1996). A contrasting view is that it depends on a general inferential mechanism, which is the same for all sorts of reasoning. One such approach postulates that individuals possess a mental logic made up from formal rules of inference (e.g., Braine & O'Brien, 1998; Rips, 1994). The central idea of these theories is that individuals construct a formal proof that a conclusion follows from the premises. Each step in the proof depends on a formal rule of inference. Errors may occur because people fail to apply a formal rule correctly (Rips, 1994, p. 153). Errors should therefore be more likely with proofs that call for a greater number of steps, or for more complex and varied steps (Rips, 1994, p. 386). As Rips (1994, p. 322) argued, individuals can make deontic inferences that do not depend on familiarity with the domain, such as:

(19) It is obligatory that P given Q.
Therefore, it is permissible that P given Q.

He suggested that such inferences can be handled in his system by the addition of modal operators akin to those proposed by logicians (see also Osherson, 1974–6). He proposed that Cheng & Holyoak's schemas (1985, 1989) could be deduction rules defined over these modal operators. However, as he pointed out, the extension of his system to account for deontic reasoning would entail more than just adding a few rules (Rips, 1994, p. 336). We therefore turn to an alternative theory that also relies on a general mechanism for reasoning.

2.3. The theory of mental models

The mental model theory postulates that reasoning depends on the meaning of assertions and on general knowledge. They are used to construct mental models of the possibilities under description (Johnson-Laird & Byrne, 1991; Polk & Newell, 1995). Three main assumptions distinguish mental models from other sorts of proposed mental representation, such as logical form and semantic networks (see, e.g., Johnson-Laird & Byrne, 2002). The first assumption is that each mental model represents a separate possibility. Hence, an exclusive disjunction, such as:

(20) Either there is a circle or else there is not a triangle

refers to two alternative possibilities. Individuals list these two possibilities when they are asked to state what is possible given the disjunction. Its meaning accordingly yields two mental models. We represent mental models using special diagrams, and we describe the notation here. Each proposed mental model is laid out in a separate row of a diagram. Hence, the two mental models of the preceding disjunction 20 are as follows:

(21) o
$\neg\Delta$

Here, 'o,' denotes a mental model of a possibility in which there is a circle, '¬' denotes negation, and so '¬Δ' denotes a model of the possibility in which there is not a triangle. Sometimes a mental model has no explicit content, but merely represents an alternative possibility whose content individuals have yet to make explicit. Diagrams denote such 'implicit' models using an ellipsis:

(22) ...

Thus, the mental models of the conditional assertion:

(23) If there is a circle then there isn't a triangle

are as follows:

(24) o $\neg\Delta$
...

The ellipsis represents an implicit model of another possibility (in which the first clause in the conditional is false). Mental models often fail to represent the status of each proposition in each model. Hence, we use the following notation:

(25) o Δ
o

In such cases, the absence of a token that occurs in another model is treated as its negation, i.e., the second model here represents a circle without a triangle. Individuals can keep track of the status of models, e.g., whether they represent real, hypothetical, or imaginary states. We use additional notation in our diagrams to indicate the status of models, e.g., we italicize those components of models that represent what is not permissible.

The theory's second assumption, which is known as the principle of 'truth,' is that mental models represent what is true according to the premises, but by default not what is false. This principle applies at two levels. At one level, models represent only the possibilities that are true given an assertion, as do the models in 21 of the disjunctive assertion 20. At a lower level, however, a model represents a clause in the premises only when the clause is true in the possibility. For example, the first model in 21 represents that there is a circle, but it does not represent explicitly that in this possibility it is *false* that there is not a triangle, i.e., there *is* a triangle. Skeptics sometimes argue that the principle of truth is nothing more than the principle that individuals represent what is mentioned in the premises. But, to represent merely what is mentioned in the premise would lead to the same representation for propositions containing different sentential connectives, such as: *If A then B*, *A or else B*, and *A and B*. In fact, the principle goes beyond mere mention to represent clauses only when they are true according to the relevant sentential connectives.

The principle of truth postulates that individuals by default do not represent what is false. But, there are exceptions that overrule the default. Individuals make 'mental footnotes' about the falsity of clauses, and if they retain these footnotes they can flesh out *mental* models into *fully explicit* models, which also represent clauses when they are false. (We forego the notation for footnotes in this paper.) The following fully explicit models, for example, represent the earlier exclusive disjunction 20:

(26) $\quad$ o $\quad \Delta$
$\neg$o $\neg\Delta$

Nevertheless, the principle of truth is the norm (see, e.g., Barres & Johnson-Laird, 2003).

The third assumption of the theory is that reasoning depends on mental models, and follows the rational principle that a conclusion is valid if it holds in all the possibilities compatible with the premises, i.e., it has no counterexamples (Johnson-Laird & Byrne, 1991). If a conclusion holds in a proportion of models, its probability is equal to that proportion, granted that the models represent equiprobable alternatives (Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni, 1999). If a conclusion holds in at least one model, it is possible given the premises (Bell & Johnson-Laird, 1998). And if it holds in none of the models, it is impossible given the premises. The theory therefore aims to unify deductive reasoning about necessity, probability, and possibility. A suite of computer programs implements the theory, and they show how inferences can be made using both mental models and fully explicit models (e.g., Bara, Bucciarelli, & Lombardo, 2001; Johnson-Laird & Savary, 1999). Likewise, a variety of evidence has corroborated the theory's predictions (see, e.g., Byrne, 1997; Johnson-Laird & Byrne, 1991, 2002).

3. The meaning and mental representation of deontic assertions

There are many frameworks in which deontic states are central—systems of morality, religion, laws, kinship, social conventions, games, and so on. In all these systems, some states are permissible and some are impermissible. Like deontic logicians (e.g., Lewis, 1973, p. 96), we will say little about the origins of these states. Of course, many states of the world are deontically neutral: they are neither permissible nor impermissible, e.g., the weather. They fall outside the scope of deontics, which concerns the actions and inactions of responsible agents. Utterances can create a deontic relation, but they can also merely describe a deontic relation. An assertion such as, 'You must leave,' can do either job depending on context. If it is spoken by someone in authority over you, then it can create an obligation. If it is spoken by a friend reminding you of an obligation, then it is merely a description. As we pointed out earlier, there is a crucial difference between the two. Violations of what is permissible according to a deontic obligation do not render the assertion false, but render the violator in breach of the obligation. Violations of a description, however, render the description false.

Our aim in this part of the paper is to formulate a theory of the meaning and mental representation of deontic assertions. The theory postulates that they refer to possible and permissible states (see Bell & Johnson-Laird, 1998, for a defense of possible states as primitives). This assumption may seem circular, but it is no more circular than postulating 'possible worlds' as underpinning the meaning of possibility (see Part 1). Alternative theories of deontic meanings could be based on moral preferences to do one action rather than another (David Over, personal communication), on sets of obligatory and non-obligatory states, on a mixture of both sorts of states, or on still other concepts (Beller, 2001). In our view, however, permissible states are the most plausible psychological foundation. A deontic relation calls for a temporal constraint: an antecedent that creates a deontic state cannot occur after the deontic state it brings about. It is prior, or at least contemporaneous, with the deontic state. But, this temporal constraint allows that an antecedent may concern current knowledge of a future action, e.g., your retirement at the age of 65 obligates you to pay into the pension plan now. The theory's first assumption is accordingly:

(I) *The principle of permissible states*. The meaning of an assertion of the form: *B is permissible* is that B is a proposition that holds in at least one member of the set of permissible states. Likewise, if the proposition, A, is possible, then the meaning of an assertion of the form: *Given A*, *B is permissible*, is that there is a joint occurrence of possible states (i.e., A's) and possible states that are also permissible (i.e., B's):

A	B
A	¬B
¬A	B
¬A	¬B

where A is possible and B is possible and permissible, and A precedes or is contemporaneous with B. In a strong sense, the assertion can mean that A is the only state in which B is permissible:

A	B
A	¬B
¬A	¬B

where we follow the convention that unless otherwise specified the first token on each line represents a possibility and the second token represents a state that is in addition permissible.

In the weak sense, A is merely one state that makes B permissible, and so when A does not occur some other state can make B permissible. In the strong sense, if A does not occur, then only not-B is permissible. There is accordingly a state in which B is not permissible:

(27) ¬A *B*

The italics in the notation here, as we mentioned earlier, denote what is *impermissible*. An example of this strong sense is:

(28) If you've finished your work then you may leave.

It refers to the following states:

(29)	Finished	Leave
	Finished	¬Leave
	¬Finished	¬Leave

where 'Finished' denotes your having finished, and 'Leave' denotes the permissible state of your leaving. What is not permissible is to leave in the state in which you have not finished. Whether or not you have finished is a matter of fact, and so that part of each state concerns what is possible. In contrast, whether or not you leave is a deontic matter, and so the corresponding part of each state refers in addition to what is permissible.

An analogous postulate governs statements of obligation:

(II) *The principle of obligatory states*. The meaning of an assertion of the form: *B is obligatory* is that the proposition B holds in all members of the set of states that are permissible, and the meaning of an assertion of the form: *Given A*, *B is obligatory*, is in its weak sense that there is a joint occurrence of states:

A	B
¬A	B
¬A	¬B

where A is possible, B is possible and permissible, and A precedes or is contemporaneous with B. In its strong sense, the assertion means that A is the only state in which B is permissible (and obligatory):

A	B
¬A	¬B

As an example, the assertion:

(30) Trespassing obligates you to leave

has the weak interpretation compatible with the following states in which leaving is permissible:

(31)	Trespass	Leave
	¬Trespass	Leave
	¬Trespass	¬Leave

In other words, if you trespass, the only permissible action is to leave, i.e., the action is obligatory, but there are other states when you are not trespassing in which you can leave. The strong sense of obligation occurs if trespassing is the unique state in which the action is obligatory.

A deontic assertion of the form, *A prohibits B*, such as: 'Your promise to stay prohibits you from leaving,' is synonymous with *A obligates not-B*. The assertion also has a strong interpretation in which A is the unique state obligating and permitting not-B. Likewise, a deontic assertion of the form, *A permits not-B*, has both a weak interpretation compatible with all four states, and a strong interpretation. Table 1 summarizes the weak and strong meanings of the four deontic relations. They can, of course, be expressed in many other ways.

The theory's principal claim is that this account exhausts the meanings of deontic relations, that is, nothing more underlies their meaning. A corollary of the postulates of permissible and obligatory states is that, as in deontic logics, the basic deontic concepts are interdefinable. The following equivalences indicate that the sentences have an interpretation in common:

Table 1
The semantics of four basic deontic relations in both weak interpretations (on the left) and strong interpretations (on the right): '+' indicates a viable interpretation

The four sorts of assertion	The permissible states given certain possibilities						
	A B A ¬B ¬A B ¬A ¬B	A B A ¬B ¬A ¬B	A B ¬A B ¬A ¬B	A B ¬A ¬B ¬A ¬B	A B ¬A ¬B	 A ¬B ¬A B ¬A ¬B	 A ¬B ¬A B
A permits B	+	+					
A permits not-B	+		+				
A obligates B				+	+		
A prohibits B						+	+

A and ¬A are possible states; B and ¬B are possible and permissible states.

(32)	**A permits B.**	≡A does not obligate not B.	≡A does not prohibit B.
	A does not permit not B.	**≡A obligates B.**	≡A prohibits not B.
	A does not permit B.	≡A obligates not B.	**≡A prohibits B.**
	A permits not B.	≡A does not obligate B	≡A does not prohibit not B.

All these relations follow from the meaning of negation and of the respective quantifications over permissible states in the principles of permissible and obligatory states. But, there are two caveats. First, assertions of the form *A permits B* are open to the weak interpretation compatible with all four states, and the paraphrases with *obligates* and *prohibits* fail on this interpretation. Second, the negative assertions are open to more than one interpretation. The salient interpretation of *A does not permit B* is that *A obligates not-B* (as shown above), which is the interpretation that follows from negating the quantifier in the principle of permissible states. But, another interpretation merely denies that 'permission' is the appropriate relation, as in: Burglary doesn't permit the police to arrest you; it obligates them to do so.

We have so far explained what sorts of states deontic assertions refer to, and we have represented these states in fully explicit models. Individuals, however, are likely to use mental models to represent these states, because the processing capacity of working memory makes it hard to hold in mind all the different fully explicit models. In the case of factual assertions, this notion is captured in the principle of truth (see Part 2). An analogous principle holds for deontic assertions: mental models represent only the salient states.

An assertion of the form *A permits B* makes salient that given A, B is permissible, and so it has the following mental models in which B is permissible:

(33) A B
A
...

The first, and most accessible, model represents the state in which A and B occur, the second model represents the state in which A occurs in the absence of B, and third model denoted by the ellipsis is an implicit model, i.e., it has no explicit content but represents that there are other permissible states (see Part 2 for a description of these conventions). If A is the unique state that renders B permissible, then there is a mental footnote on the implicit model to indicate that B is exhaustively represented in the explicit model. In principle, reasoners who retain the footnote can flesh out the mental models into *fully explicit* models corresponding to the states in Table 1. Without the footnote, however, there is nothing to constrain the fleshing out of the models, and so the fully explicit models represent all four states. As a corollary, only if A is the unique state rendering B permissible, can individuals construct a model in which B is *impermissible*:

(34) ¬A *B*

As we stated earlier, we use italics to represent what is impermissible. An assertion of the form, *A permits not-B*, has the same mental models as those above except that ¬ B occurs in place of B.

An assertion of the form *A obligates B* has the following mental models:

(35) A B
. . .

where there is a mental footnote that the implicit model cannot contain instances of A. As a corollary, individuals can construct a model representing that given A, not-B is impermissible:

(36) A ¬*B*

Unlike the previous assertions, those of the form *A prohibits B* have as their mental model one that represents what is *not* permissible:

(37) A *B*

As before, the italics denote that B is not permissible (cf. the developmental evidence reviewed in Part 2). If, say, you are prohibited from smoking then what you must *not* do is to smoke. As a corollary, the mental models of what *is* permissible are as follows:

(38) A ¬ B
. . .

where there is a mental footnote that the implicit model cannot contain instances of not-A. The permissible models can therefore be fleshed out into the fully explicit models shown in Table 1.

We summarize these mental models of deontic assertions in the following assumption, which is analogous to the principle of truth:

(III) *The principle of deontic mental models*. The mental models of *A permits B* represent that given A, B is permissible; the mental models of *A obligates B* represent that given A, only B is permissible; and the first mental model of *A prohibits B* represents that given A, *B* is not permissible, and the second mental model represents that given A, not-B is permissible.

One consequence of this principle, as we will see, is that compelling but fallacious inferences should occur in reasoning from deontic premises, i.e., so-called 'illusory

Table 2
The mental models for two ways of expressing each of the four basic deontic relations

Deontic relations	Mental models	
	Permissible	Impermissible
1. A permits B:	A B A ...	
1′. Not-A prohibits B:	 ¬A ¬B ...	¬A *B*
2. A permits not-B:	A ¬B A ...	
2′. Not-A obligates B:	¬A B ...	
3. A obligates B:	A B ...	
3′. Not-A permits not-B:	¬A ¬B ¬A ...	
4. A prohibits B:	 A ¬B ...	A *B*
4′. Not-A permits B:	¬A B ¬A ...	

Only *prohibits* yields a mental model of what is impermissible, and this model is more salient than the mental models of what is permissible.

inferences.' Another consequence is that different ways of expressing a given deontic relation have different mental models, though the same fully explicit models. The choice between two deontic expressions with an interpretation in common is likely to reflect a speaker's communicative intentions. Table 2 presents the mental models of two ways of expressing each of the four main deontic relations. For example, as the table shows, an assertion of the form, *A prohibits B*, makes the impermissible state salient, whereas an assertion of the form, *A obligates not-B*, makes a permissible state salient. Hence, a speaker is likely to choose whichever form of assertion reflects what the listener should focus on.

4. A test of the theory of meaning and representation

4.1. Experiment 1: Models of deontic relations

Our first experiment was designed to test the theory of the meaning of deontic assertions and of their mental representation in mental models. We asked 24 student volunteers at the University of Turin to list what was permissible and what was impermissible for each of the two ways of expressing the basic deontic relations in Table 2. The theory makes two main predictions. First, the participants' listings should correspond to the fully explicit models in Table 1. Second, they should start their listings with the situations corresponding to the mental models in Table 2. As a corollary, the two different ways of expressing the same deontic relation should yield similar overall listings, but the assertions based on *permits* should be more likely to yield interpretations in which all four states are permissible (the weak interpretation of the relation).

4.1.1. Method

The participants had to list what was permissible (è consentito) and what was impermissible (non è consentito) for the eight sorts of deontic assertion in Table 2. Each sort of assertion occurred with two different contents for a total of 16 trials in a random order. The contents were assigned randomly to the set of problems so that in the experiment as a whole each content occurred equally often with each of the eight sorts of assertion, and each content occurred twice within each set of materials but with a different deontic assertion. The contents concerned everyday matters that were appropriate for all of the eight sorts of assertion, e.g., as translated from the Italian:

(39) Tax-payers who support charities are obligated to claim a rebate on their taxes.

The full set of contents is presented in the Appendix. The participants, who were tested individually, were told that the experiment concerned how people understand certain sorts of sentence. The key instructions were as follows: 'Your task is to consider each assertion to be true and then to envisage the states that are permissible and those that are not permissible given the truth of the assertion.' The experimenter then explained that the participants could list the states in any order, and that there were four states to consider for each assertion. The four states were illustrated with the assertion, 'Bosses who are lazy force employees to work hard.' If, as rarely happened, the participants failed to list one of the four states during the experiment, then the experimenter pointed out the omission.

4.1.2. Results and discussion

Table 3 presents the percentages of the main interpretations of the eight sorts of assertion. These interpretations corroborated the theory's predictions. The majority of interpretations lie on the diagonal of the table, but, as the theory predicted, any assertion based on *permits* can yield the weak interpretation in which all four states are permissible (the first column in the table). All 24 participants conformed to the predictions more often than not. In particular, for the four assertions based on *permits*, the participants listed all four states as permissible more often than for their paraphrases with other verbs (Wilcoxon test, $z = 3.4$, $p < .001$). The bias towards the weak interpretation of *permits* might be more pronounced than the everyday interpretation of the verb. Daily usage may be more informative, and individuals are not normally called on to list all the permissible and impermissible states corresponding to a deontic assertion. Similarly, the everyday interpretation of *obligates*

Table 3
The percentages of the most frequent interpretations in Experiment 1 ($n = 24$)

The eight sorts of assertion	The permissible states in the participants' interpretations						
	A B A ¬B ¬A B ¬A ¬B	A B A ¬B ¬A ¬B	A B A ¬B ¬A B	A B ¬A B ¬A ¬B	A B ¬A ¬B	A ¬B ¬A B ¬A ¬B	A ¬B ¬A B
1. A permits B	52	33					
1′. Not-A prohibits B		85					
2. A permits not-B	65		17				
2′. Not-A obligates B			88				10
3. A obligates B				90	10		
3′. Not-A permits not-B	46			38			
4. A prohibits B						83	6
4′. Not-A permits B	40					44	

The balances of the percentages were interpretations occurring on less than 10% of trials.

and *prohibits* may be modulated by knowledge that rules out the state unique to their weak interpretation. As a consequence, these relations may receive the strong interpretation more often than occurred in the present experiment. What is important, however, is that the participants made all the predicted interpretations, seldom departed from them, and showed a general consensus about the meanings of deontic assertions.

Table 4 presents the percentages of trials on which the participants started each listing with the state corresponding to the salient mental model of the assertion. They

Table 4
The eight deontic relations and the percentages of states listed first in Experiment 1

Deontic relations	Percentages of states listed first	
1. A permits B:	A B	100
1′. Not-A prohibits B:	¬A *B*	60
	¬A ¬B	25
	A B	15
2. A obligates B:	A B	100
2′. Not-A permits not-B:	¬A ¬B	42
	¬A B	54
3. A prohibits B:	A *B*	83
	A ¬B	13
3′. Not-A permits B:	¬A B	85
	A *B*	13
4. A permits not-B:	A ¬B	35
	A B	61
4′. Not-A obligates B:	¬A B	88
	A B	12

The balances of the percentages are states listed on less than 5% of trials. An italicized *B* indicates that the state was listed as impermissible; the other states were listed as permissible.

did so on 74% of the trials, and all 24 participants started with this state more often than not. As the table shows, there were no marked differences in the interpretations from one participant to another. In particular, the interpretations of the affirmative assertions were quite uniform. Likewise, the assertions based on *prohibits* tended to elicit first the impermissible state (21 out of the 24 participants listed the impermissible state first on one or both occasions). The assertions containing negation, however, yielded a more varied choice of initial instance. For instance, with assertions of the form *A permits not-B*, only 35% of trials began with the state, A and not-B, and many participants listed first the state in which A and B occurred. After the participants had listed their first case for an assertion, they did not use a fixed strategy, but varied in their listings. Some continued with permissible states, but some switched to an impermissible state and then back to a permissible state, and so on. In sum, the results corroborated the model theory. We consider alternative accounts in the General discussion.

5. Deductive reasoning from deontic assertions

How do individuals make deontic inferences? For example, how would they draw the following conclusion:

(40) Having children obligates you to look after them. (A obligates B.)
To look after children prohibits you from leaving them unattended. (B prohibits C.)
Therefore, having children prohibits you from leaving them unattended. (A prohibits C.)

The inference is valid, i.e., if the premises are true, then the conclusion must be true based on the interpretations of *obligates* and *prohibits* in Table 1, and on the judgements of naïve individuals shown in Table 3. The model theory makes the following assumption:

(IV) *The principle of deontic reasoning*. Deontic deductions are based on mental models rather than fully explicit models of assertions.

As we will see later, this assumption predicts systematic errors in reasoning. But, we need first to explain how models are conjoined in order to reason. The principles are the same as those for conjoining the possibilities compatible with sentential connectives (see, e.g., Johnson-Laird & Savary, 1999). They are summarized in Table 5 below.

The first principle of interpretation (rule 0 in Table 5) concerns the conjunction of mental models in which one model represents a state that is not represented in the other model, e.g., the conjunction of a model of A and B with a model of B. If A occurs in some other model from the set from which B is drawn, then, as we have mentioned, its absence counts as negation. But, otherwise, its absence is equivalent to affirmation. We now consider two contrasting inferences to illustrate the rules for conjoining models. The first inference is example 40 above with premises of the form:

Table 5
The principles for forming a conjunction of two sets of mental models or fully explicit models of permissible states

0. Consider the conjunction of a pair of mental models in which a state, A, in one model is not represented in the other model, e.g.:
   A B and B
   If the other models from which B is drawn contain an occurrence of A, then its absence is treated as negation; otherwise its absence is treated as an affirmation. The rules for conjoining models are:
1. The conjunction of a pair of models that represent, respectively, a state and its negation yields the null model (akin to the empty set):
   A B and ¬A B yield nil
   A contradiction also occurs if B is permissible in one model but impermissible in the other model. But, when B is possible in one model, but permissible (or impermissible) in the other model, the resulting conjunction inherits this deontic status.
2. The conjunction of any model with the null model yields the null model:
   A B and nil yield nil
3. The conjunction of two models with explicit content conjoins them and drops any duplicates:
   A B and B C yield A B C
4. The conjunction of an explicit model with an implicit model yields the null model by default:
   A B and . . . yield nil
   But, if none of the atomic propositions in the explicit model (A B) is represented in the set of models containing the implicit model, then the conjunction yields the explicit model, e.g.:
   A B and . . . yield A B
5. The conjunction of two implicit models yields an implicit model:
   . . . and . . . yield . . .

The procedures apply to each pairwise conjunction of models from the two sets. Each procedure is presented with an accompanying example. Only principles 1 through 3 apply to the conjunction of fully explicit models.

(41) A obligates B.
B prohibits C.

The first premise has the following mental models in which B is permissible (as shown in Table 2):

(42) A B
. . .

The second premise has the following mental model in which C is impermissible:

(43) B *C*

To integrate the two sets of models, each model of the first premise has to be conjoined with the model of the second premise, i.e., the pairs have to be formed into conjunctions. A pair of models can be directly conjoined unless they refer to inconsistent states, in which case the result is the null model akin to the empty set (rule 1 in Table 5). Likewise, the conjunction of the null model with any model also yields the null model (rule 2 in Table 5). In the example, the first model of the first premise and the model of the second premise are consistent, and they can be directly conjoined (using rule 3 in Table 5) to yield the following model in which C is impermissible:

(44) A B *C*

The conjunction of the implicit model (represented by the ellipsis) with the explicit model of the second premise is also the null model. The reason is that one element in the explicit model, B, occurs elsewhere in the set containing the implicit model, i.e., it occurs in the first model of the first premise. This occurrence implies that the implicit model includes the case of ¬B, which is inconsistent with B. The conjunction of the two sets of models is accordingly the single model:

(45) A B *C*

This model corresponds to the conclusion: *A prohibits C*, and so, according to the theory, naïve individuals should draw this conclusion. The fully explicit models of the premises (see Table 1) support the same conclusion.

The contrasting inference is one in which the verbs in the previous example are swapped from one premise to the other:

(46) A prohibits B.
B obligates C.

The mental model of the first premise is:

(47) A *B*

where B is impermissible. The models of the second premise are:

(48) B C
...

where B is possible. The conjunction of the two explicit models yields the following model (according to rule 3):

(49) A B *C*

from which it follows that *A prohibits C*. But, reasoners may consider instead what is permissible given the first premise, namely:

(50) A ¬B
...

The task now is to conjoin these models with those of the second premise:

(51) B C
...

The two initial models in each set represent, respectively, a situation, B, and its negation, ¬B, and so their conjunction yields the null model (rule 1). The combination of the explicit model: A ¬B, with the implicit model in the second set, yields the explicit model: A ¬B (rule 4). In this case, unlike the first inference, neither A nor ¬B occur in any model of the second premise. The conjunction of the implicit model in the first set with the explicit model, B C, in the second set likewise yields: B C. Finally, the conjunction of the implicit model in the first set with the implicit model in the second set yields an implicit model (rule 5 in Table 5). The overall conjunction of the two sets of models is accordingly:

(52) A ¬B
B C
...

Because C does not occur in the model representing A, these models correspond to the conclusion:

(53) A permits not-C

or to its equivalents, such as:

(54) A does not obligate C.

These conclusions are valid, whereas the initial conclusion derived above, that *A prohibits C*, is invalid. The conjunction of the fully explicit models of the two premises (see Table 1) yields the following models:

(55) A ¬B C
A ¬B ¬C
¬A B C
¬A ¬B C
¬A ¬B ¬C

These models are based on the weak interpretations of the two verbs, but even with a strong interpretation of either verb, the first state remains, i.e., given A, C is permissible. Hence, A does not prohibit C. The fully explicit models also show that the premises yield all four states interrelating A and C. Hence, any of the following four conclusions in their weak senses are valid: *A permits C*, *A permits not-C*, *Not-A permits C*, and *Not-A permits not-C*. We have now illustrated the principles for conjoining sets of models, and Table 5 summarizes them with illustrative examples.

5.1. Experiment 2: Reasoning from pairs of deontic premises

The principle of deontic reasoning asserts that individuals should rely on mental models in making deontic inferences. Three predictions follow. First, reasoners should tend to draw those conclusions corresponding to mental models. Second, inferences depending on one mental model should be easier than those depending on multiple mental models. Third, premises yielding more than one mental model with an explicit content should elicit a greater variety of responses than premises yielding only one mental model with an explicit content. As a corollary, the premises with multiple models should also elicit more 'no valid conclusion' responses than the latter premises, because individuals should have difficulty in finding a conclusion that holds over more than one model. The present experiment was designed to test these three predictions. The experiment used all 16 possible pairs of premises based on the four main deontic relations in the arrangement exemplified here:

(56) People who work in big companies are obligated to go on holiday in August.
People who go on holiday in August are not obligated to stay in big hotels.
What follows?

We used premises of the form *Not-A obligates B* in place of those of the form, *A permits not-B*, in order to avoid the weak interpretation compatible with all four states. The 16 forms of inference are shown in Table 6 together with their mental models and the conclusions that these models yield. Eight of the problems yield strong valid conclusions, and eight of the problems yield only weak valid conclusions, i.e., their fully explicit models include all four states relating A and C.

5.1.1. Method

We tested 20 student volunteers at the University of Turin. They had to state what, if anything, followed from each of the 16 pairs of premises presented in a different random order. The contents of the problems concerned everyday matters, and we made two different random assignments of the materials to the 16 forms of inference. Half the participants were tested with one assignment, and half the participants were tested with the other assignment. The participants were told that the experiment was about reasoning. The key instructions were as follows: 'Your task is to read carefully the premises, and to try to draw a conclusion. To draw a conclusion, you must relate the terms that are not directly related in the premises.' They were told that there might not be any conclusion that had to be true given the premises, and in which case they should say so. The participants carried out one practice trial. They were allowed to take as much time as they needed to make their responses.

5.1.2. Results and discussion

Table 7 presents the numbers of participants making the predicted responses for each problem and those unpredicted responses that occurred more than once to a given problem. 91% of the responses were predicted, and the mean number of predicted responses per problem was 18.1 out of 20. Only three predicted responses were not made by any participant. All the participants made more predicted responses than unpredicted responses. Likewise, every problem yielded more predicted responses than unpredicted responses. As Table 7 shows, the problems with one mental model tended to yield unanimity in the conclusions that the participants drew (the mean amount of statistical information in the distribution of responses was 0.38 bits), whereas the problems with multiple models yielded a variety of responses (the mean amount of statistical information in the distribution of responses was 1.23 bits). This predicted difference in the distributions of responses was reliable (Mann-Whitney $W = 23.0$, $z = 2.15$, $p < .025$). Likewise, as predicted, the response 'no valid conclusion' occurred more often with multiple-model problems (31% per problem) than with one-model problems (2% per problem; Mann-Whitney $W = 19.5$, $z = 2.55$, $p < .01$). There were three multiple-model problems that were exceptions and that elicited few or none of the predicted 'no valid conclusion'

Table 6
The 16 inferences in Experiment 2, their mental models, and the conclusions that they yield

The second premise	The first premise			
	A obligates B	A permits B	A prohibits B	Not-A obligates B
B obligates C	A B C . . . A OBLIGATES C	A B C A . . . A PERMITS C nvc	A ¬B B C . . . A PERMITS NOT-C nvc A *B* *C* A prohibits C	¬A B C . . . NOT-A OBLIGATES C
B permits C	A B C A B . . . A PERMITS C nvc	A B C A B A . . . A PERMITS C nvc	A ¬B B C B . . . A permits not-C nvc A *B* *C* A *B* A PROHIBITS C	¬A B C ¬A B . . . NOT-A PERMITS C nvc
B prohibits C	A B ¬C . . . A PROHIBITS C	A B ¬C A . . . A PERMITS NOT-C nvc A B *C* A prohibits C	A ¬B B ¬C . . . A PERMITS C nvc A *B* *C* A prohibits C	¬A B ¬C . . . NOT-A PROHIBITS C
Not-B obligates C	A B ¬B C . . . A PERMITS NOT-C nvc	A ¬B C ¬B C A B A . . . A PERMITS NOT-C nvc	A ¬B C . . . A OBLIGATES C	¬A B ¬B C . . . NOT-A PERMITS NOT-C nvc

For premises based on *prohibits*, the table shows the models of permissible states and, where they yield different conclusions, the models of impermissible states (in italics) in the lower half of an entry. Valid conclusions are in CAPITALS and invalid conclusions are in lower case; and 'nvc' denotes the response 'no valid conclusion.' Some valid conclusions based on *permits* depend on its weak interpretation, which is compatible with all four states.

Table 7
The numbers of participants in Experiment 2 ($n = 20$) making the predicted responses, or their equivalents where shown

The second premise	The first premise							
	A obligates B		A permits B		A prohibits B		Not-A obligates B	
B obligates C	A obligates C:	20	A permits C:	10	A not obligates C:	10	Not-A obligates C:	20
			nvc:	5	nvc:	7		
			A obligates C:	5	A prohibits C:	3		
B permits C	A permits C:	20	A permits C:	18	A permits not C:	0	Not-A permits C:	18
	nvc:	0	nvc:	2	nvc:	7	nvc:	1
					A prohibits C:	12		
B prohibits C	A prohibits C:	19	A permits not-C:	0	A not prohibits C:	5	Not-A prohibits C:	16
			nvc:	9	nvc:	12	*A prohibits C*:	3
			A prohibits C:	7	A prohibits C:	3		
			A not prohibits C:	4				
Not-B obligates C	A not obligates C:	9	A permits not-C:	5	A obligates C:	17	Not-A not obligates C:	10
	nvc:	8	nvc:	10	*nvc*:	2	nvc:	7
	A obligates C:	2	*A permits C*:	4				

Responses that are not predicted by the model theory and that occurred more than once are in italics. ‘nvc’ denotes the response of ‘No valid conclusion.’

responses (see Table 7). Two of these problems had two models that differed only in the presence or absence of a token representing a single proposition. For example, the problem:

(57) A obligates B.
B permits C.

yields the mental models:

(58) A B C
A B
...

The two models differ only in the presence or absence of a token of C. Hence, it is relatively simple to draw the conclusion: *A permits C*. In sum, the results corroborate the model theory's predictions.

6. Illusory inferences in deontic reasoning

Is there any way to strengthen the claim that reasoners rely on mental models? The answer depends on an unexpected prediction of the model theory. Consider the following sort of exclusive disjunction:

(59) You are permitted to have only one of the following two actions:
You have the soup.
You have the salad.

You are allowed to have either the course described in the first assertion or else the course described in the second assertion, but not both, and so your mental representation of the disjunction should consist of two mental models of what you are permitted:

(60) Soup
Salad

Hence, if someone asked you whether you could have the soup, then you should infer that it would be permissible, because such a state is represented in your models of what is permissible. Now, consider the following more complex disjunction:

(61) Problem 1. You are permitted to carry out only one of the following actions:
Action 1: Take the apple or the orange, or both.
Action 2: Take the pear or the orange, or both.
Are you permitted to take the orange?

You are allowed to carry out either action 1 or else action 2, but not both. The mental models of action 1 represent what it is permissible to take:

(62) Apple
Orange
Apple Orange

They support the conclusion that it is permissible to take the orange. Likewise, the mental models of action 2 support the same conclusion. Hence, if reasoners rely on mental models of deontic states, then they should respond: 'Yes, I'm permitted to take the orange'. However, the response is an illusion. If you did take the orange then you would have carried out both action 1 and action 2, contrary to the rubric that you are permitted to carry out only one of them. Unlike mental models, the fully explicit models of the problem take into account that when one action is permissible the other action is impermissible. They show that only two states are permissible:

(63)	Apple	¬Pear	¬Orange
	¬Apple	Pear	¬Orange

You can take the apple or else you can take the pear, but you cannot take the orange. The correct response to the question is therefore: 'No, it is not permissible to take the orange.'

A simple control problem for the illusion has the same premises, but asks the question:

(64) Problem 1′. Are you permitted to take the pear?

The mental models of action 2 yield the answer: 'yes'; and this answer is correct. Hence, reasoners should be correct even if they fail to consider that when they carry out one action, they must not carry out the other action.

These predicted illusions arise from the principle of deontic mental models. In particular, as corroborated in Experiment 1, mental models of what is permitted represent only permissible states (see Tables 2 and 4). The aim of our next experiment was to examine whether the predicted illusions occur.

6.1. Experiment 3: Illusory inferences about permission

The experiment examined two sorts of illusion and their corresponding control inferences. One sort was an illusion of permissibility in which reasoners should infer that an action is permissible when in fact it is not permissible. The other sort was an illusion of impermissibility in which reasoners should infer that an action is impermissible when in fact it is permissible. This second sort of illusion should occur in the following problem:

(65) Problem 2. You are permitted to carry out only one of the following actions:
Take the apple or the orange, or both.
Take the pear and the orange.
Are you permitted to take the pear without the orange?

The mental models of the first assertion represent what it is permissible to take:

(66)	Apple	
		Orange
	Apple	Orange

And the mental model of the second assertion represents what it is permissible to take:

(67) Pear Orange

This model predicts that reasoners should respond: No, you cannot take the pear without taking the orange. Once again, however, the response is an illusion. If you took both the pear and the orange, then you would have carried out both actions in the problem. Hence, you are permitted to take the pear, but only if you do not take the orange, and so the correct response is: Yes. The corresponding control problem has exactly the same premises, but the more complex question:

(68) Problem 2′. Are you permitted to take neither the apple, nor the orange, nor the pear?

The mental models above yield the answer: 'No,' and this response is correct.

6.1.1. Method

Two hundred and forty students at the University of Turin voluntarily carried out four inferential problems (the two illusions and the two control problems) as part of a course. Four different contents were assigned twice to the four problems, and half the participants were tested with one assignment and half the participants were tested with the other assignment. The problems were assembled in booklets in four different random orders, and the booklets were assigned to balanced numbers of participants. The booklets contained the following instructions (translated from the Italian): "This is an experiment on how people reason, it is not a test of intelligence or a personality test. The booklet you received has a problem printed on each page. Read carefully the statements on each page and try to consider what follows, then write down your response on the page. Once you have turned a page, do not go back to any previous problems. You have no time limit."

Table 8
The four forms of problem in Experiment 3 (in abbreviation) and the percentages of correct responses to them

Illusory problems	Control problems
1. Permitted only one action: Take A or B, or both Take C or B, or both Permitted to take B?	1′. Permitted only one action: Take A or B, or both Take C or B, or both Permitted to take C?
Correct answer, No: 3	Correct answer, Yes: 94
2. Permitted only one action: Take A or B, or both Take C and B Permitted to take C without B?	2′. Permitted only one action: Take A or B, or both Take C and B Permitted to take neither A nor B nor C?
Correct answer, Yes: 10	Correct answer, No: 87

6.1.2. Results and discussion

Table 8 presents the percentages of correct responses to the four sorts of problem. There was no significant effect of content on performance, and so we pooled the results. As the model theory predicted, the participants responded much more accurately to the control problems (90% correct) than to the illusions (7% correct): 229 of the 240 participants performed better with the control problems than with the illusory problems, three performed better with the illusory problems than the controls, and the remaining eight were ties (binomial test, $p \ll .0001$).

Readers might suspect that the difficulty of the illusions pertains to the artificial nature of the task, or to a possible difficulty in understanding a rubric such as: 'You are permitted to carry out only one of the following actions.' The participants might accordingly have found the problems confusing and guessed their answers. This possibility seems unlikely because the control problems were based on exactly the same premises, differing only in their questions, and yet the participants were right about them on 90% of occasions. Moreover, the performance on the illusions was much lower than chance, and guessing should tend to yield a chance level of performance.

6.2. Experiment 4: An antidote to deontic illusions

Previous studies have demonstrated illusory inferences from descriptions based on sentential connectives (e.g., Johnson-Laird & Savary, 1999; Johnson-Laird et al., 2000), and it has been difficult to find effective antidotes to them, other than to teach the participants to consider both what is true and what is false (Yang & Johnson-Laird, 2000b). Reasoning about deontic states, however, offers the possibility of a theoretically excellent antidote, which is to couch problems in the form of *prohibitions*. As Experiment 1 showed, given a prohibition, individuals tend to think first about what is impermissible (see Table 4). A disjunction of deontic states can be expressed, as in the previous experiment, using the rubric: 'You are permitted to carry out only one of the following actions.' But, the same disjunction can be expressed in a logically equivalent way using the following prohibition: 'You are prohibited from carrying out more than one of the following actions.' The model theory predicts that in this case individuals will focus on what is prohibited, i.e., they must not carry out both actions. We can therefore construct a problem that is equivalent to problem 1 in the previous experiment, but in which the rubric refers to what is prohibited:

(69) You are prohibited from carrying out more than one of the following actions:
To take the apple or the orange, or both.
To take the pear or the orange, or both.
Are you permitted to take the orange?

Reasoners should construct the mental models of what it is *impermissible* to take:

(70) *Apple* *Pear*
Orange

As these models show, if they take the apple and the pear, then they carry out both actions; likewise, if they take the orange, then they carry out both actions. Hence, they should respond correctly that it is not permissible to take the orange. It follows that the illusory inferences should be reduced for prohibitions.

The corresponding control problem couched in terms of *prohibits* has the same premises as those above, but the question is:

(71) Are you permitted to take the pear?

The mental models of what is impermissible do not include the state of taking the pear alone, and so reasoners should make the right response, 'Yes,' but the task should be harder than a control problem based on the rubric in terms of *permits*, which yields a model that directly represents the permissibility of taking the pear. The same principles apply to illusory problem 2 and its control. In sum, the model theory predicts that the control problems should be easier than the illusions. But, it also predicts an interaction. Illusory problems should yield a better performance with *prohibits* than with *permits*, whereas control problems should yield a better performance with *permits* than with *prohibits*. The main aim of our final experiment was to investigate these predictions.

6.2.1. Method

We tested individually 40 student volunteers at the University of Turin with eight different problems consisting of the four problems from the previous experiment with the rubric based on *permits*, and the same four problems with the rubric based on *prohibits*. Eight different contents were assigned twice to the problems, and half the participants were tested with one assignment and half the participants were tested with the other assignment. Within these assignments, the problems were presented in four different random orders to the participants. The instructions were the same as those in the previous experiment.

Table 9
The total numbers of participants ($n = 40$) making correct responses to the problems in Experiment 4

Illusory problems			Control problems		
	Permitted only one action:	Prohibited from more than one:		Permitted only one action:	Prohibited from more than one:
Take A or B, or both Take C or B, or both Permitted B?			Take A or B, or both Take C or B, or both Permitted C?		
Correct: No	1	10	Correct: Yes	40	25
Take A or B, or both Take C and B Permitted C without B?			Take A or B, or both Take C and B Permitted neither A nor B nor C?		
Correct: Yes	2	16	Correct: No	27	12

6.2.2. Results and discussion

Table 9 presents the percentages of correct responses to the eight problems. There was no significant effect of content on performance, and so we pooled the results. As the model theory predicted, the participants responded more accurately to the control problems (65% correct) than to the illusions (18% correct): 33 of the 40 participants performed better with the control problems than with the illusory problems, and none performed better with the illusory problems than with the controls (binomial test, $p = .5^{33}$). The predicted interaction was also reliable: 28 out of the 40 participants did better with illusory problems based on *prohibits* but better with control problems based on *permits*, there were 10 ties, and only 2 participants yielded results contrary to the interaction (Wilcoxon test, $z = 4.7$, $p < .0001$). Designed comparisons showed that the illusory problems yielded a better performance with *prohibits* than with *permits* (Wilcoxon test, $z = 3.62$, $p < .0005$), whereas the control problems yielded a better performance with *permits* than with *prohibits* (Wilcoxon test, $z = 4.04$, $p < .0001$).

Although the verb *prohibits* alleviated the illusions, it did so at a cost of impairing performance with the control problems. A study of antidotes to illusions based on descriptions of quantified sets had a similar effect, improving performance with illusions but impairing it with controls (Yang & Johnson-Laird, 2000b).

7. General discussion

Our aim has been to formulate a theory of the meaning of the fundamental deontic assertions, of their mental representation, and of elementary deontic reasoning. The semantic theory is based on two principles:

(I) Permissible states: The meaning of *A permits B* is that in at least one state of A, B is permissible, where A precedes or is contemporaneous with B. In its strong sense, only A makes B permissible; in its weak sense, all four contingencies can occur (see Table 1).

(II) Obligatory states: The meaning of *A obligates B* is that in all states of A, B is permissible and not-B is impermissible. In its strong sense, B is not permissible in any other state; in its weak sense, B is permissible without A (see Table 1).

Experiment 1 corroborated that naïve individuals enumerated these proposed sets in a reliable way (see Table 3). The theory further postulates that individuals normally represent deontic assertions in mental models rather than in fully explicit models:

(III) Deontic mental models: The mental models of deontic assertions represent salient states, that is, of what is permissible for *A permits B*, *A permits not-B*, and *A obligates B*, but of what is impermissible for *A prohibits B* (see Table 2).

Experiment 1 also corroborated this principle, because the participants tended to list first the states corresponding to these mental models (see Table 4).

To explain reasoning from deontic assertions, the theory postulates:

(IV) Deontic reasoning: Deontic inferences should tend to be based on mental models rather than fully explicit models.

The principles for conjoining models are summarized in Table 5. Given, say, premises of the form:

(72) A obligates B.
B prohibits C.

the mental models of the premises yield the conclusion (see Table 6):

(73) A prohibits C.

Experiment 2 corroborated the principle. In general, the participants tended to draw those conclusions supported by the mental models of the premises, drawing one and the same conclusion for problems with a single explicit mental model, but making the various predicted responses for problems with more than one explicit mental model (see Table 7).

The principle of deontic reasoning predicts that individuals should be susceptible to illusory inferences about what they are allowed to do. Consider, for example:

(74) You are permitted to carry out only one of the following actions:
Take the apple or the orange, or both.
Take the pear or the orange, or both.
Are you permitted to take the orange?

Reasoners should think about what they can take if they carry out the first action. There are three possibilities, depending on whether they take the apple or the orange, or both, and so they should infer that they are allowed to take the orange. They should draw the same conclusion from the second assertion. But, if they were to take the orange, then they would carry out actions in accordance with both assertions, contrary to the rubric of the problem. Mental models merely represent what is permissible, and only fully explicit models embody the consequences of the impermissibility of the other action.

Experiment 3 established that individuals do succumb to such illusions (see Table 8). Illusions in reasoning about descriptions have been demonstrated in previous studies (e.g., Johnson-Laird & Savary, 1999), but the present experiment established a novel phenomenon. Individuals draw illusory inferences about what they are allowed to do. One reaction to this phenomenon is that they go wrong because the task is unusual, unnatural, or artificial. As a critic suggested, the participants will assume that they are dealing with normal discourse, not subtly worded trick problems. The criticism seems plausible, but it overlooks several phenomena. First, the participants do very well with the control problems, which are based on the same premises, and merely ask a different question (see Table 8). The same principle of deontic reasoning predicts both inaccurate performance with the illusory inferences and accurate performance with the control inferences. Second, this principle also predicts performance in the Experiment 2, which used quite ordinary wording of problems. Third, as Experiment 3 showed, it is possible to reduce the illusions in a predictable way. Assertions of prohibition, as Experiment 1 confirmed, tend to make salient what is *not* permissible. It follows that the rubric:

(75) You are prohibited from carrying out more than one of the following actions

should lead reasoners to infer what is impermissible. If this rubric is combined with the preceding problem with its question about the orange, then individuals should be likely to infer that they are not allowed to take the orange. Experiment 4 corroborated such a reduction of the illusions, but, as the theory also predicted, there was a concomitant increase in the difficulty of the control problems with the rubric couched in terms of prohibition as opposed to permission (see Table 9).

Alternative theories of deontic reasoning have been based on formal rules of inference (e.g., Braine & O'Brien, 1998; Osherson, 1974–6; Rips, 1994), pragmatic schemas (Cheng & Holyoak, 1985; Holyoak & Cheng, 1995), and innate modules for reasoning about special cases (Cosmides, 1989; Cummins, 1996). Most of these theories, like most empirical studies of deontic reasoning, have focussed on the selection task, which we reviewed in Part 1. None of them deals with the deontic inferences that we have examined, but perhaps they could be extended to make predictions about these inferences. Our results present three challenges to putative extensions of these theories. They need to account, first, for the meanings of modal terms (including their strong and weak interpretations), second, for the salience of some permissible situations over others, and, third, for the occurrence of illusory inferences. This last problem is severe for theories based on rules. Valid rules cannot yield illusions; invalid rules run the risk of internal inconsistency.

The model theory meets these challenges. It postulates that the foundation of human rationality is the ability to appreciate the force of counterexamples. They show that arguments are invalid (Johnson-Laird & Byrne, 1991). Irrational performance arises from a failure to flesh out mental models into fully explicit models. Mental models usually suffice to yield valid conclusions, but a few inferences in the set of possible inferences yield errors if reasoners rely solely on mental models. In the case of factual discourse, mental models are based on the principle of truth, and so they do not represent explicitly what is false (see Part 2). In the case of deontic discourse, mental models are based on the principle of deontic mental models, and so they make salient what is permissible in the case of permissions and obligations, but what is impermissible in the case of prohibitions. The principles of truth and deontic models are special cases of a superordinate constraint. When a complete representation—a set of fully explicit models—would overwhelm the limited processing capacity of working memory, the reasoning system relies instead on the most useful subset—what is true in the case of factual discourse, and what is permissible or impermissible depending on the deontic assertion.

Is deontic reasoning special? That is, does it depend on mental processes that are different from those underlying other sorts of reasoning? Some psychologists, as we have seen, argue for special mechanisms for deontic reasoning. In contrast, the model theory postulates that what differs is, not the mechanism, but the contents on which it operates. The meanings of deontic assertions refer to sets of permissible and impermissible states. Reasoning, however, depends on the same general processes that apply elsewhere. Reasoners construct a set of mental models for the premises and draw a conclusion, if any, which the models support. As we pointed out in the introduction, deontics raises many psychological problems beyond the meanings, representations, and implications of deontic assertions. Whether the present theory can contribute to their solution is an open question.

Acknowledgments

This research was made possible in part by a grant to the first author from the M.I.U.R. of Italy, FIRB project (research code: RBAUOIJEYW_001), and in part by a grant to the second author from the National Science Foundation (Grant BCS-0076287) to study strategies in reasoning. We thank Bruno Bara, Ruth Byrne, Jonathan Evans, Tom Ormerod, Keith Stenning, and an anonymous reviewer, for their helpful criticisms of earlier drafts of this paper. And we thank especially Earl Hunt for his comments and editing much beyond the call of duty.

Appendix. Lexical contents used in Experiment 1

1. Tax-payers who support charities are permitted/permitted not/obligated/forbidden to claim a rebate on their taxes.
2. Bankers who make credits are permitted/permitted not/obligated/forbidden to increase the interest rates.
3. Invalids who use wheelchairs are permitted/permitted not/obligated/forbidden to use small elevators.
4. Musicians who play wind instruments are permitted/permitted not/obligated/forbidden to arrive one hour before the concert.
5. Students who attend prestigious universities are permitted/permitted not/obligated/ forbidden to wear shorts.
6. The competitors in the singing competition are permitted/permitted not/obligated/forbidden to participate in the musical competition.
7. People who work in big companies are permitted/not permitted/obligated/forbidden to go on holiday in August.
8. The nurses who do the day-work are permitted/permitted not/obligated/forbidden to attend the adjournment course.

References

Almor, A., & Sloman, S. A. (1996). Is deontic reasoning special? *Psychological Review, 103*, 374–380.

Bara, B. G., Bucciarelli, M., & Lombardo, V. (2001). Model theory of deduction: A unified computational approach. *Cognitive Science, 25*, 839–901.

Barres, P., & Johnson-Laird, P. N. (2003). On imagining what is true (and what is false). *Thinking & Reasoning, 9*, 1–42.

Beller, S. (2001). A model theory of social norms. In J. D. Moore & K. Stenning (Eds.), *Proceeding of the twenty third annual conference of the cognitive science society* (pp. 63–68). Mahwah, NJ: Erlbaum.

Bell, V., & Johnson-Laird, P. N. (1998). A model theory of modal reasoning. *Cognitive Science, 22*, 25–51.

Braine, M. D. S., & O'Brien, D. P. (Eds.). (1998). *Mental logic*. Mahwah, NJ: Erlbaum.

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Is there an innate module for deontic reasoning. In J. García-Madruga, N. Carriedo, & M. J. González-Labra (Eds.), *Mental models in reasoning* (pp. 227–239). Madrid: UNED.

Byrne, R. M. J. (1997). Cognitive processes in counterfactual thinking about what might have been. In D. Medin (Ed.). *The psychology of learning and motivation, advances in research and theory* (Vol. 37, pp. 105–154). San Diego, CA: Academic Press.

Byrnes, J. P., & Duff, M. A. (1989). Young children's comprehension of modal expressions. *Cognitive Development, 4*, 369–387.

Cheng, P. N., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 391–416.

Cheng, P. N., & Holyoak, K. J. (1989). On the natural selection of reasoning theories. *Cognition, 33*, 285–313.

Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*, 187–276.

Cummins, D. D. (1996). Evidence of deontic reasoning in 3- and 4-year-old children. *Memory & Cognition, 24*, 823–829.

Day, C. (1996). Understanding of the French modal verbs 'pouvoir' and 'devoir' in school children and adults. *Current Psychology of Cognition, 15*, 535–553.

Day, C., & Caron, J. (1991). Le developpment de la comprehension des verbes modaux 'pouvoir' et 'devoir' chez des enfants de six a treize ans. *Archives de Psychologie, 59*, 55–69.

Evans, J. St. B. T., & Twyman-Musgrove, J. (1998). Conditional reasoning with inducements and advice. *Cognition, 69*, 811–816.

Goldvarg, Y., & Johnson-Laird, P. N. (2001). Naïve causality: A mental model theory of causal meaning and reasoning. *Cognitive Science, 25*, 565–610.

Gralinski, H. J., & Kopp, C. B. (1993). Everyday rules for behavior: mothers' requests to young children. *Developmental Psychology, 29*, 573–584.

Green, D. W., McClelland, A., Muckli, L., & Simmons, C. (1999). Arguments and deontic decisions. *Acta Psychologica, 101*, 27–47.

Hilpinen, R. (Ed.). (1971). *Deontic logic: Introductory and systematic readings*. Dordrecht, Holland: Reidel.

Hirst, W., & Weil, J. (1982). Acquisition of the epistemic and deontic meaning of modals. *Journal of Child Language, 9*, 659–666.

Holyoak, K., & Cheng, P. (1995). Pragmatic reasoning with a point of view. *Thinking & Reasoning, 1*, 289–314.

Hughes, G. E., & Cresswell, M. (1984). *A companion to modal logic*. New York: Methuen.

Johnson-Laird, P. N. (1978). The meaning of modality. *Cognitive Science, 2*, 17–26.

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

Johnson-Laird, P. N., & Byrne, R. M. J. (1996). A model point of view: A comment on Holyoak and Cheng. *Thinking & Reasoning, 1*, 339–350.

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., Legrenzi, P., Girotto, V., Legrenzi, M. S., & Caverni, J. P. (1999). Naive probability: A model theory of extensional reasoning. *Psychological Review, 106*, 62–88.

Johnson-Laird, P. N., Legrenzi, P., Girotto, P., & Legrenzi, M. S. (2000). Illusions in reasoning about consistency. *Science, 288*, 531–532.

Johnson-Laird, P. N., & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition, 71*, 191–229.

Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development, 69*, 706–720.

Klaczynski, P. A., & Narasimhani, G. (1998). Representations as mediators of adolescents deductive reasoning. *Developmental Psychology, 34*, 865–881.

Kohlberg, L. (1986). A current statement of some theoretical issues. In S. Modgil & C. Modgil (Eds.), *Lawrence Kohlberg: Consensus and controversy* (pp. 485–546). Philadelphia: Falmer.

Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, D. (2000). Semantic analyses for dyadic deontic logic. In D. Lewis (Ed.), *Papers in ethics and social philosophy* (pp. 5–19). Cambridge: Cambridge University Press (Originally published 1974).

Lyons, J. (1977). *Semantics* (Vol. 3). Cambridge: Cambridge University Press.

Manktelow, K. I., & Over, D. E. (1990). *Inference and understanding*. Routledge.

Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition, 39*, 85–105.

Manktelow, K. I., & Over, D. E. (1995). Deontic reasoning. In S. E. Newstead & J. St. B. T. Evans (Eds.), *Perspectives on thinking and reasoning: Essays in honour of Peter Wason* (pp. 91–114). Hillsdale, NJ: Erlbaum.

Noveck, I. A., Ho, S., & Sera, M. (1996). Children's understanding of epistemic modals. *Journal of Child Language, 23*, 621–643.

Osherson, D. N. (1974–6). *Logical abilities in children* (Vols. 1–4). Hillsdale, NJ: Lawrence Erlbaum Associates.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review, 102*, 533–566.

Quelhas, A. C., & Byrne, R. M. J. (2003). Reasoning with deontic counterfactual conditionals. *Thinking & Reasoning*, in press.

Ray, J. L., Reynolds, R. A., & Carranza, E. (1989). Understanding choice utterances. *Quarterly journal of experimental psychology: Human experimental psychology A, 41*, 829–848.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Schaeken, W., & d'Ydewalle, G. (1996). Must reasoning. *Acta Psychologica, 92*, 209–220.

Searle, J. R. (1969). *Speech acts*. Cambridge: Cambridge University Press.

Shepard, S. C. (1982). From deontic to epistemic: An analysis of modals in the history of English, creoles, and language acquisition. In A. Ahlquist (Ed.), *Papers from the fifth international conference on historical linguistics* (pp. 232–316). Amsterdam: Benjamins.

Staller, A., & Petta, P. (2001). Introducing emotions into the computational study of social norms: A first evaluation. *Journal of Artificial Societies and Social Simulation, 4*(1), U27–U60.

Steedman, M. J. (1977). Verbs, time, and modality. *Cognitive Science, 1*, 216–234.

Stenning, K., & van Lambalgen, M. (2004). A little logic goes a long way: Basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science, 28*, 481–529.

van Fraasen, B. (1973). The logic of conditional obligation. *Journal of Philosophical Logic, 1*, 417–438.

Walker, L. J. (1988). The development of moral reasoning. *Annals of Child Development, 5*, 33–78.

Wason, P. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmondsworth, UK: Penguin.

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

Yang, Y., & Johnson-Laird, P. N. (2000b). How to eliminate illusions in quantified reasoning. *Memory & Cognition, 28*, 1050–1059.