Mem Cogn
DOI 10.3758/s13421-012-0188-2

The consistency of disjunctive assertions

P. N. Johnson-Laird · Max Lotstein · Ruth M. J. Byrne

Abstract In two experiments, we established a new phenomenon in reasoning from disjunctions of the grammatical form *either A or else B*, where A and B are clauses. When individuals have to assess whether pairs of assertions can be true at the same time, they tend to focus on the truth of each clause of an exclusive disjunction (and ignore the concurrent falsity of the other clause). Hence, they succumb to illusions of consistency and of inconsistency with pairs consisting of a disjunction and a conjunction (Experiment 1), and with simpler problems consisting of pairs of disjunctions, such as *eIther there is a pie or else there is a cake* and *Either there isn't a pie or else there is a cake* (Experiment 2), that appear to be consistent with one another, but in fact are not. These results corroborate the theory that reasoning depends on envisaging models of possibilities.

Keywords Consistency · Disjunction · Logic · Mental models · Inference · Reasoning

When individuals think intuitively about the world, they tend to have in mind just one possibility. They can make most sorts of inference in this way, including many based on conditional assertions, such as *If the President is in Baku, then the Secretary is too* (Byrne, 2005; Byrne & Johnson-Laird, 2009). In contrast, disjunctions—that is, states of affairs naturally described using "or"—force people to think about alternative possibilities. Being in a disjunctive state of mind leads to difficulty in reasoning: for instance,

> The President is in Baku or the Secretary is.
> In fact, the President is not in Baku.
> Therefore, the Secretary is.

The inference is *valid*; that is, the conclusion is true in every case in which the premises are true (Jeffrey, 1981, p. 1). Moreover, the inference is valid whether the disjunction is *inclusive*, which allows that both its clauses could be true (*A or B, or both*), or *exclusive*, which disallows that both its clauses could be true (*A or else B, but not both*). In a test of members of the general public, only 30% drew the valid conclusion from the disjunctive inference, whereas 48% drew valid conclusions from analogous premises based on a conditional assertion, *If A then B*, and the categorical denial *not B* (Johnson-Laird, Byrne, & Schaeken, 1992).

The standard view of disjunction is that it has an inclusive meaning, unless the rider "but not both" is added to its assertion (Barrett & Stenner, 1971; Chierchia & McConnell-Ginet, 1990; Crain & Khlentzos, 2007; Grice, 1975; Kamp & Reyle, 1993, pp. 191–192). Proponents of the standard view claim that without this rider, few disjunctions in daily life are exclusive—granted that such an interpretation demands that when both disjuncts are true, the disjunction as a whole is false. Both inclusive and exclusive disjunctions can be used to convey various speech acts, such as threats, promises, and inducements (Fillenbaum, 1974). But no definitive list of speech acts exists, so we assume that these interpretations are a consequence of the pragmatics of disjunctions rather than an element of their underlying meaning (see Miller & Johnson-Laird, 1976, sec. 7.4.4). The standard view accordingly allows for conversational implicatures (Grice, 1975), but an implicature to an exclusive interpretation will take time to interpret (Noveck,

P. N. Johnson-Laird · M. Lotstein
Princeton University,
Princeton, NJ, USA

R. M. J. Byrne (✉)
School of Psychology and Institute of Neuroscience, Trinity
College Dublin, University of Dublin,
Dublin 2, Ireland
e-mail: rmbyrne@tcd.ie

Published online: 01 March 2012

Chierchia, Chevaux, Guelminger, & Sylvestre, 2002). An exclusive interpretation can indeed take longer to evaluate as true or false (Chevallier et al., 2008), but it has been shown that inferences from exclusive disjunctions are easier to make than those from inclusive disjunctions (Bauer & Johnson-Laird, 1993; Johnson-Laird et al., 1992). The standard view of disjunction is embodied in theories of reasoning based on formal rules of inference (e.g., Braine & O'Brien, 1998; Macnamara, 1986; Rips, 1994), which include rules couched only for inclusive disjunctions and treat exclusive disjunctions as equivalent to assertions of the form *A or B, and not both A and B*.

In contrast to the standard view, the theory of mental models explains—in fact, predicts—the difficulty of inferences from disjunctions (Johnson-Laird & Byrne, 1991). Individuals use lexical meanings, grammatical structure, and knowledge to construct models of the possibilities to which assertions refer. A conclusion is inferred to be valid if it holds in all of the resulting models of possibilities. Mental models are iconic, in that each part of a model corresponds to each part of what it represents (see Hartshorne, Weiss, & Burks, 1933). An iconic representation of a disjunction accordingly calls for separate models of each possibility to which the disjunction refers. Hence, the disjunction *The President is in Baku or the Secretary is* calls for a model in which the President is in Baku, a model in which the Secretary is in Baku, and, given an inclusive interpretation, a model in which both of them are in Baku. We summarize these models in the following diagram representing who is in Baku, where each row denotes a mental model of a distinct possibility:

President	
	Secretary
President	Secretary

Disjunctions are difficult, so the theory avers, because of the processing load of two or three distinct mental models on working memory. One obvious prediction is that inferences should be easier if the disjunction is exclusive, because it then refers to only the first two of the models in the preceding diagram. This prediction is correct: Inferences from exclusive disjunctions *are* easier than those from inclusive disjunctions (Bauer & Johnson-Laird, 1993; Johnson-Laird et al., 1992). This result could be reconciled with formal rules of inference, provided that rules were included for exclusive disjunctions, but it is hard to see how such rules could *predict* the phenomenon, and to the best of our knowledge, no one has framed such a theory.

Experimental evidence has shown that individuals tend to interpret "or else" as signaling an exclusive interpretation. For instance, when participants listed what was possible given assertions of the forms *(A or else B) and C* and *(A or else B) or else C*, they listed the possibilities corresponding to the mental models of exclusive disjunctions; for instance, 82% of their interpretations of the latter assertion listed three separate possibilities: A, B, and C (Barres & Johnson-Laird, 2003). As a consequence, "or else" is odd, and even unacceptable, if the contents of a disjunction imply an inclusive interpretation, as in

> He was alive yesterday or else today.

The implication is that if he wasn't alive yesterday, then he was alive today, which is bizarre. As Zimmermann (2000, p. 269) wrote, in a linguistic analysis: "[or else] does involve disjointness, if we assume that *else* restricts the second alternative to the cases in which the first one is out." Hence, it makes better sense to drop "else" from our example: *He was alive yesterday or today.* The exclusive interpretation of "or else" is also borne out in corpora of English. For example, a corpus of 425 million words in recent American English (http://corpus.byu.edu/coca/) yields the following uses of "or else" in assertions, which we found on the first page of over 2,000 examples:

> They have to succeed, or else this thing is going to get a lot worse.
> He must have run through a fortune, or else he has some sort of a system.
> I will find Aaron or else leave word for him that I am waiting.

These assertions all support inferences from the affirmation of the first clause to the negation of the second clause:

> They succeeded, so this thing isn't going to get a lot worse.
> He did run through a fortune, so he doesn't have some sort of a system.
> I found Aaron, so therefore I won't leave word for him that I am waiting.

All of these inferences are valid only in the case of an exclusive interpretation. A corollary is that "or else" is often used in threats, such as

> Don't do that or else I'll punish you,

with the force that there are two possibilities: In one, the listener carries out the action and is open to punishment, and in the other, the listener refrains from the action and avoids punishment (Fillenbaum, 1974). In sum, "or else" is common in everyday discourse, and it tends to imply an exclusive interpretation.

The theory of mental models—henceforth, the "model theory"—also illuminates a phenomenon that appears at first sight to support the standard theory (D. Over, personal communication). Inferences of the following sort are intuitively valid: *If A or B then C; A and B*, therefore *C*. For instance,

> If the car has been stolen or damaged in an accident, then the insurance company will pay up.

The car has been stolen and damaged in an accident.
Therefore, the insurance company will pay up.

It seems impossible to interpret "or" in the first premise as exclusive, but, according to the standard theory, this interpretation should be feasible if there were an exclusive sense of "or" in English. The model theory, however, predicts that the interpretation of sentential connectives can be modulated both by the contents of the clauses that they connect and by general knowledge (Johnson-Laird & Byrne, 2002), and that this modulation is not canceled as a result of inserting a disjunction into the "if" clause of a conditional. Here is an example in which an inference of the same sort as before is *not* intuitively valid:

If it has all the sexual characteristics of a female or it has all the sexual characteristics of a male, then it is not a hermaphrodite.
It has the sexual characteristics of a female and it has the sexual characteristics of a male.
Therefore, it is not a hermaphrodite.

Indeed, a simple recipe generates an indefinite number of invalid conclusions C from premises of the form *If A or B then C; A and B*: The disjunct A should imply C, the disjunct B should imply C, but A and B should imply not-C: for instance, *If he's married to Ann or he's married to Beth, then he's a monogamist.* Modulation of this sort explains the well-known effects of contents and context on the interpretation of simple disjunctions (Gualmini, Crain, & Meroni, 2000; Newstead & Griggs, 1983; Newstead, Griggs, & Chrostowski, 1984; Roberge, 1976a, 1976b, 1977, 1978). Meaning and knowledge can modulate the interpretation of "or," blocking the construction of a model of a possibility or adding temporal and other relations between models of disjuncts (Juhos, Quelhas, & Johnson-Laird, 2012).

Mental models, such as those in our earlier diagram, are based on a principle of *truth*: They represent only those situations that are possible given an assertion, and each model of a possibility represents only what is true in each possibility according to the assertion (see, e.g., Johnson-Laird & Savary, 1999). However, when the intellectual demands of a task are not so great—as in, say, listing possibilities—individuals can flesh out mental models into *fully explicit* models, which represent both what is true and what is false with respect to the clauses in the premises. The fully explicit models of the inclusive disjunction *The President is in Baku or the Secretary is* are shown in the following diagram, where each row represents a model of who is and who isn't in Baku:

President	not-Secretary
not-President	Secretary
President	Secretary

where "not" is a symbol denoting negation, which is used here to represent a false affirmative. Reasoning is a demanding task, so individuals tend to rely on mental models rather than fully explicit models. As a result, they are liable to make systematic fallacies in those inferences for which falsity plays a critical role (see, e.g., Goldvarg & Johnson-Laird, 2000; Khemlani & Johnson-Laird, 2009; Yang & Johnson-Laird, 2000a, 2000b). The following example gives rise to such a fallacy:

Only one of the following statements is true:
At least some of the plastic beads are not red.
None of the plastic beads are red.
Is it possible that none of the red beads are plastic?

The mental models represent the truth of the first premise and, separately, the truth of the second premise. In the second case, the model in which none of the plastic beads are red yields the conclusion that none of the red beads are plastic. Hence, the theory predicts that individuals should respond "yes" to the question. And they do. However, this response is an illusion, because according to the initial statement, when the second assertion is true, the first assertion is false, and its falsity implies that all of the plastic beads are red. Hence, the correct response to the question is "no," it is impossible that none of the red beads are plastic (see Yang & Johnson-Laird, 2000a, 2000b). In contrast, individuals tend to respond correctly to similar control problems. Other illusions depend on conditionals (e.g., Johnson-Laird & Savary, 1999), but the semantics of conditionals is a highly controversial matter (e.g., Rips, 1997; Stenning & van Lambalgen, 2008), so it is important that illusory inferences occur with assertions other than conditionals (e.g., Khemlani & Johnson-Laird, 2009; Walsh & Johnson-Laird, 2004). The challenge to the model theory is to demonstrate illusory inferences that cannot be explained by the standard theory and that call for a form of reasoning in which the participants do not have to draw any conclusion, let alone one based on conditionals. In the present article, we report new fallacies of this sort.

The task that we investigated calls for individuals to assess whether or not a set of assertions is consistent. The assessment of consistency is intimately related to logical reasoning, even though it seems quite different. The maintenance of consistent beliefs is a hallmark of rationality, and its importance goes back to Aristotle (Edghill, 1928). Indeed, the intimate relation between consistency and deduction is central to many systems of logic. In these systems, the validity of an inference is established by showing that the negation of its conclusion conjoined with the premises yields an inconsistent set of assertions (see, e.g., Jeffrey, 1981). In other words, valid deduction and the assessment of consistency are merely different sides of deductive competence. Moreover, the task is within the competence of participants who have not mastered logic, provided that it is couched in these terms: "Could all of these assertions be true at one and the same time?" (see, e.g., Johnson-Laird, Girotto, & Legrenzi, 2004; Legrenzi, Girotto, & Johnson-Laird, 2003). An affirmative answer implies that there is a possibility in which all of the assertions are true, so it follows that the assertions are consistent one with another. Likewise, a negative answer implies that the assertions are not consistent one with another.

The task of judging consistency poses a challenge to current theories of deductive reasoning, such as accounts based on formal rules of inference (e.g., Braine & O'Brien, 1998; Rips, 1994), on suppositions (e.g., Evans & Over, 2004), and on probabilistic considerations (e.g., Geiger & Oberauer, 2010; Oaksford & Chater, 2007). These theories can represent that assertions are inconsistent, but at present they offer no account of how individuals reach this evaluation or of which problems are more likely to yield errors. Such extensions may be forthcoming. Meanwhile, however, the present study focuses on the predictions of the model theory and avoids trying to pit the model theory against these other accounts, because of uncertainty about their predictions for the evaluation of consistency.

Experiment 1: Disjunctions and illusions of consistency

The model theory's principle of truth predicts that illusions should occur in the evaluation of consistency. As an example, consider this problem, which we used in the experiment:

> Either there is an apple on the table or else there isn't both a banana and a cherry.
> There is an apple and not a banana on the table.
> Can both of these assertions be true at the same time?

Table 1 summarizes the processes governing the construction of the models for disjunctive assertion in this case. As it shows, the mental models of what is on the table according to this assertion are as follows:

apple		
	not-banana	not-cherry
	banana	not-cherry
	not-banana	cherry

The decision about consistency is simple: Both assertions can be true at the same time if, and only if, there is at least one possibility in which they both hold. The second assertion is true in the first of the preceding models, because there is an apple but not a banana, so the theory predicts that reasoners should respond "yes," the two assertions can be true at the same time. (The absence of an entity from a model implies its negation if the entity occurs in another model in the same set.) As Table 1 also shows, the fully explicit models of the disjunctive assertion take into account that when its first clause is true, its second clause is false, and vice versa. Hence, the fully explicit models are as follows:

apple	banana	cherry
not-apple	not-banana	not-cherry
not-apple	banana	not-cherry
not-apple	not-banana	cherry

These models show that the correct answer to the question above is "no": An apple without a banana on the table is inconsistent with the first assertion.

The contrast between mental models and fully explicit models motivates three other sorts of problem in which a given disjunction is paired with different second assertions. The second assertion

> There isn't a banana and there isn't a cherry.

holds in the second of both the mental and fully explicit models. Hence, reasoners should respond "yes," and in this case the response is correct. This sort of problem is accordingly a control for the illusion of consistency. The second assertion

> There is an apple, a banana, and a cherry.

does not occur in any mental model of the disjunction. Hence, reasoners should respond "no" (the two assertions cannot be true at the same time). However, this assertion does hold in the first of the fully explicit models, so problems of this sort should give rise to an illusion of inconsistency: Reasoners infer that the two assertions cannot both be true, but in fact they can be. A control for this sort of illusion is the second assertion

> There isn't an apple, but there is a banana and a cherry.

The mental models predict a "no" response, and in this case it is correct, because the second assertion doesn't hold in the fully explicit models, either. The theory accordingly predicts the occurrence of two sorts of illusions: illusions of consistency, in which individuals infer that assertions are consistent when in fact they are inconsistent, and illusions of inconsistency, in which individuals infer that assertions are inconsistent when in fact they are consistent. For four different sorts of disjunction, in this experiment we compared performance for the two sorts of illusions with performance for their respective control problems, which should yield the correct responses.

Method

Design, participants, and materials The experiment was carried out online, and 19 participants on Amazon's Mechanical Turk acted as their own controls and carried out all of the different sorts of problems. There were four main sorts of problem, which we list here with their abbrevations, in which the first letter states the predicted response and the

Table 1 Construction of mental models and of fully explicit models of disjunctions

0. The mental models of the two sorts of disjunction are as follows:

A or B, or both:			*A or B, but not both*:	
	A			A
		B			B
	A	B		

The process of constructing models for an assertion, such as: *Either there is an apple on the table or else there isn't both a banana and a cherry* (*A or else not both B and C*) is as follows.

1. The model of a conjunction, *B and C*, is: B C.

2. Negation in *not both B and C* calls for recovering the complement of the unnegated models. There are four possible models based on *B* and *C*, and the complement of B C is therefore the three models:

not-B	not-C
B	not-C
not-B	C

2. The mental models of an exclusive disjunction, *or else*, consist of the union of the models of the two disjuncts. Hence, *A or else not both B and C* has these mental models:

A		
	not-B	not-C
	B	not-C
	not-B	C

The absence of an entity, such as B or C from the first of these models, implies its negation, not-B and not-C, because these entities occur in other models in the set.

3. Fully explicit models take into account that when one disjunct is true, the other disjunct is false:

A	B	C
not-A	not-B	not-C
not-A	B	not-C
not-A	not-B	C

second letter states the correct response ("C" for consistent and "I" for inconsistent): illusions of consistency (C/I), their controls (C/C), illusions of inconsistency (I/C), and their controls (I/I). These four sorts of problems were based on each of the four sorts of disjunction:

Either A, or else not both B and C.
Not both A and B, or else C.
Either A, or else A and B and C.
Either A and B and C, or else C.

Table 2 below presents the resulting 16 problems, their mental models according to the principle of truth, and their fully explicit models. The contents of the problems consisted of assertions about three fruits—*apples*, *bananas*, and *cherries*—and for each participant, a computer program assigned these fruits randomly to the problems, which were also presented to each participant in a different random order.

Procedure The instructions explained the general nature of the task, and their crux was "On each problem, you'll be presented with two sentences and you will be asked to determine whether both could be true at the same time. In the space provided, please begin your answer with either 'Yes' or 'No,' followed by an explanation of your reasons." Prior to the experiment proper, the participants carried out a simple practice problem and were told that their explanations should be complete sentences. We paid them a small fee, and 1 participant was chosen at random to receive a bonus. We took the usual precautions for online studies: We sampled only participants from North America and filtered out those who were not native speakers of English, who had no experience of online experiments, who had negligible latencies, who took more than 5 min on any one problem, or who made the same response for all of the problems.

Results and discussion

The results corroborated the model theory's predictions at a highly significant level. Table 2 presents the results for the

Table 2 The 16 forms of problem in the four categories in Experiment 1, their statuses, the mental models and fully explicit models of the disjunctions, and the percentages of correct responses

	Problems		% Correct responses	Mental models of the first assertions	Fully explicit models of the first assertions
	First assertion: Either A, or else not both B and C			a ¬b c b ¬c ¬b ¬c*	a b c ¬a ¬b c ¬a ¬b ¬c ¬a b ¬c
1	Not-B and not-C	(C/C)	79		
2	A and not-B	(C/I)	32		
3	Not-A and B and C	(I /I)	74		
4	A and B and C	(I/C)	53		
	First assertion: Not both A and B, or else C			¬a b a ¬b ¬a ¬b c	¬a b ¬c a ¬b ¬c ¬a ¬b ¬c a b c
5	Not-A and not-B	(C/C)	84		
6	Not-B and C	(C/I)	32		
7	A and B and not-C	(I/I)	58		
8	A and B and C	(I/C)	47		
	First assertion: Either A, or else A and B and C			a a b c	a ¬b c a b¬ c a ¬b¬ c
9	A and B	(C/C)	63		
10	B and C	(C/I)	32		
11	Not A and not B and not C	(I/I)	89		
12	A and B and not C	(I/C)	26		
	First assertion: Either A and B and C, or else C			a b c c	¬a b c a ¬b c ¬a ¬b c
13	B and C	(C/C)	58		
14	A and B	(C/I)	53		
15	Not A and not B and not C	(I/I)	79		
16	A and not B	(I/C)	5		

"C" stands for "consistent," "I" stands for inconsistent; the first letter is the predicted response, and the second is the correct response. "¬" denotes "not." *In problems of this sort, individuals often construct only the first and last of these mental models (Khemlani, Orenes, & Johnson-Laird, 2012), but this result has no bearing on the predictions in the present experiment.

individual problems. The overall percentages of correct responses were as follows:

Controls for consistency	(C/C):	71%
Illusions of consistency	(C/I):	37%
Controls for inconsistency	(I/I):	75%
Illusions of inconsistency	(I/C):	33%

A total of 13 of the 19 participants were more accurate on the controls for consistency than on the illusions of consistency, and the remaining 6 were ties (binomial test of 13 out of 13, $p < .000125$); in addition, 15 of the 19 participants were more accurate on the controls for inconsistency than on the illusions of inconsistency, and there was 1 tie (binomial test of 15 out of 18, $p < .005$). As Table 2 shows, the problems within each of the four sorts varied in difficulty. With Problem 14, a closer examination of the participants' responses suggested that, although performance was around chance, the essence of the illusion occurred universally; that is, the participants failed to recognize the impossibility of the first disjunct. The evidence that we cited in the introduction showed that "or else" tended to be interpreted as an exclusive disjunction. We could not use "not both" to signal exclusivity in the present experiment, because it would have confused the participants in assertions of the sort *A or else not both B and C*. However, previous results with similar assertions showed that individuals listed possibilities consistent with the exclusive interpretation (see Barres & Johnson-Laird, 2003). Moreover, an inclusive interpretation would not change the status of two of the illusions of inconsistency (Problems 12 and 16 in Table 2), and the participants did succumb to them, if anything to a greater degree than to the other illusions.

Experiment 2: The simplest disjunctive illusions

The problems in the previous experiment concerned pairs of assertions referring to three entities. The aim of the present experiment was to test whether illusions would occur with simpler problems that concerned pairs of disjunctions referring to only two entities. The reader was invited to consider whether both of the following assertions could be true at the same time:

> Either the pie is on the table or else the cake is on the table.
> Either the pie isn't on the table or else the cake is on the table.

The mental models of the first disjunction represent two possibilities for what is on the table:

pie	
	cake

Likewise, the mental models of the second disjunction are

not-pie	
	cake

Hence, the two assertions both hold in the possibility in which the cake is on the table, so it seems that the two disjunctions are consistent with one another. But this evaluation is an illusion that occurs because mental models fail to represent what is false. The fully explicit models of the first disjunction are:

pie	not-cake
not-pie	cake

and the fully explicit models of the second disjunction are

not-pie	not-cake
pie	cake

As they show, the two assertions do not refer to a single possibility in common, and so it is impossible for both assertions to be true at the same time. In contrast, consider the following pair of assertions:

> Either the pie is on the table or else the cake is on the table.
> Either the pie isn't on the table or else the cake isn't on the table.

The mental models of the first assertion (shown above) allow that either the pie isn't on the table (the second model) or else the cake isn't on the table (the first model), so reasoners should conclude that the two assertions could be true at the same time. This response is correct, because the fully explicit models of the second assertion are identical to those for the first assertion. In this experiment, we examined a set of such illusory inferences and their respective controls.

Method

Design and materials The experiment was carried out online, and 22 participants on Amazon's Mechanical Turk acted as their own controls and carried out two different instances of the four sorts of problem: illusions of consistency and inconsistency, and their respective controls—that is, a total of eight problems. In Table 3 we present the eight sorts of problems, their mental models, and their fully explicit models. The contents of the problems consisted of assertions about the presence on a table of various common objects with monosyllabic names—for instance, *the pie*, *the cake*, and *the tart*. For each problem, two of the three objects were randomly selected. For each participant, these contents were assigned at random to each sort of problem and the problems were presented in a different random order.

Procedure The instructions were the same as those in the previous experiment, which called for the participants to decide whether or not both sentences in a pair could be true at the same time. As in the previous experiment, we paid the participants a small fee, and 1 participant was chosen at random to receive a bonus. We took the usual precautions for online studies (described in the previous experiment).

Results

Overall, the results corroborated the model theory's predictions. Table 3 presents the results for the individual problems. The overall percentages of correct responses were as follows:

Controls for consistency:	(C/C)	77%
Illusions of consistency:	(C/I)	43%
Controls for inconsistency:	(I/I)	91%
Illusions of inconsistency:	(I/C)	41%

A total of 14 of the 22 participants were more accurate on the controls for consistency than on the illusions of consistency, and 6 of the remainder were ties (binomial test, 14 out of 16, $p < .0025$); in addition, 16 of the 21 participants were more accurate on the controls for inconsistency than on the illusions of inconsistency, and there were 5 ties (binomial test, 16 out of 16, $p < .000025$). The difference between performance on the controls and illusions of consistency appeared to be smaller than that for performance on the

Table 3 The eight forms of problem in the four categories in Experiment 2, their statuses, the mental models and fully explicit models of the disjunctions, and the percentages of correct responses

Problems		% Correct Responses	Mental models of the first assertions	Fully explicit models of the first assertions
First assertion: Either A or else B.			a b	a ¬b ¬a b
1 Either not-A or else not-B.	(C/C)	75		
2 Either not-A or else B.	(C/I)	23		
First assertion: A or B, or else B.			a a b b	a ¬b
3 A and not-B.	(C/C)	81		
4 A and B.	(C/I)	64		
First assertion: Either A and B, or else not-B.			a b ¬b	a b a ¬b ¬a ¬b
5 Not-A and B.	(I/I)	95		
6 A and not-B.	(I/C)	62		
First assertion: Either A, or else not-A and B.			a ¬a b	a b a ¬b ¬a b
7 Not-A and not-B.	(I/I)	86		
8 A and B	(I/C)	20		

"C" stands for "consistent," "I" stands for inconsistent; the first letter is the predicted response, and the second is the correct response. "¬" denotes "not."

controls and illusions of inconsistency, but the interaction was not reliable (Wilcoxon test, z = 0.96, p > .3). One anomaly was that illusory Problems 4 and 6 elicited better performance than did the other illusions, and control Problem 5 elicited outstanding performance. As their explanations showed, some participants simplified the first assertion to an exclusive disjunction: *either A or else B* for Problem 4, so they inferred correctly the inconsistency of *A and B*. Likewise, some participants simplified the first assertion to an exclusive disjunction: *either A or else not-B* for Problems 5 and 6, and so they made the correct response to these problems. As in the previous experiment, the interpretation of "or else" as inclusive cannot explain the results, because the two illusions of inconsistency remain illusions in comparison with their controls (Problems 6 and 8 in Table 3).

General discussion

Assertions based on "or" present a greater challenge to reasoning than do those based on any other sentential connective, because reasoners are forced to deliberate about alternative possibilities. Evidence corroborates this claim; for instance, inferences from *if* and *and* are easier to draw than inferences from *or* (García-Madruga, Moreno, Carriedo, Gutiérrez, & Johnson-Laird, 2001; Johnson-Laird et al., 1992). Dual-process theories of reasoning distinguish between rapid intuitive inferences and slow deliberative inferences (see, e.g., Evans, 2003; Kahneman, 2011; Stanovich, 1999; Verschueren, Schaeken, & d'Ydewalle, 2005). The model theory postulates that the intuitive system can cope with only one model at a time (Johnson-Laird, 1983, 2006), so one cause of the difficulty of disjunctive reasoning is that it is beyond the scope of simple intuitions: It normally depends on multiple possibilities, and these possibilities even challenge the deliberative system. We therefore defend a theory based on mental models, which predicts the phenomena of disjunctive reasoning: Exclusive disjunctions elicit only two models of possibilities, so they yield easier inferences than do inclusive disjunctions, which elicit three models of possibilities (Bauer & Johnson-Laird, 1993; Johnson-Laird et al., 1992).

A central postulate of the model theory is that individuals focus on what is true at the expense of what is false. A corollary is that the mental models of an exclusive disjunction of the grammatical form *Either A or else not both B and C* represent a possibility in which *A* holds, along with the possibilities corresponding to the negated conjunction. It follows that individuals should infer that this assertion could be true at the same time as one of the form *A and not-B*, because this assertion holds in the model of *A* alone being true. They are mistaken, because when *A* holds, *not both B and C* is false, so all three propositions hold in this possibility: A B C. Experiment 1 showed that such fallacies occur, whereas individuals do quite well with control problems. Experiment 2 corroborated these findings using pairs of assertions such as

> Either there is a pie on the table or else there is a cake on the table.
> Either there isn't a pie on the table or else there is a cake on the table.

Both assertions seem to be true if there is a cake on the table. However, once again, this inference is erroneous. According to the first disjunction, if there is a cake on the table, then there isn't a pie, but according to the second disjunction, if there is a cake on the table, then there is a pie. Hence, there can't be a cake on the table. In fact, the two assertions do not refer to any possibilities in common. Even simple pairs of disjunctions can be misleading, as the model theory predicts.

The standard theory of disjunction, as we outlined in the introduction, takes inclusive disjunction as basic, so it cannot predict that inferences from exclusive disjunctions are easier than those from inclusive disjunctions. The proponents of this theory might argue that the present experiments are artificial, but as we also showed in the introduction, assertions containing "or else" do occur in daily life, and they do tend to be interpreted as exclusive disjunctions. Moreover, such an argument fails to explain the satisfactory level of performance with the control inferences. Alternatives to the model theory include accounts that rely on formal rules of inference (e.g., Braine & O'Brien, 1998; Rips, 1994; Stenning & van Lambalgen, 2008), on suppositions (e.g., Evans & Over, 2004), and on probabilistic heuristics (e.g., Geiger & Oberauer, 2010; Oaksford & Chater, 2010). At present, none of these theories gives an account of how reasoners evaluate the consistency of sets of assertions. In our view, it would be premature to use our results to advance arguments against these theories. Instead, we have shown that the model theory predicts the results, so our findings present a challenge to the alternative theories. The proponents of the various alternative theories might consider how to extend their theories to provide at least a post hoc explanation of the phenomena. A step in this direction has already been taken, in the efforts to integrate probabilistic considerations with mental models (e.g., Geiger & Oberauer, 2010; Oaksford & Chater, 2010). The real challenge is to extend such analyses to disjunctions and to the task, not of inferring conclusions, but of assessing consistency.

Naïve reasoners know the conditions in which both inclusive and exclusive disjunctions are true (Barres & Johnson-Laird, 2003), but they often have difficulty putting this knowledge to work in inferring from the falsity of one clause that the other clause must hold. In addition, the present experiments corroborated a further consequence of the theory: When individuals have to infer whether pairs of assertions can both be true at the same time, they tend to focus on the truth of disjuncts at the expense of their falsity. Certain pairs of assertions therefore lead to systematic errors in performance in comparison with control pairs. Disjunctions are difficult, and their difficulty arises because it is hard to have sensible intuitions about them. To reason correctly, individuals have to hold in mind models of more than one possibility, to represent both what is true and what is false, and to think about these possibilities deliberatively. In other words, they must rely not on a single mental model of a possibility, but on fully explicit models of a set of possibilities. And such reasoning is difficult for most of us, most of the time.

Author note This research was supported by Grant SES 0844851 to the first author from the National Science Foundation to study deductive and probabilistic reasoning, and by a grant to the third author from the Arts and Social Sciences Benefactions Fund of Trinity College Dublin. The authors are grateful to Geoffrey Goodwin, Eoin Gubbins, Sangeet Khemlani, Niklas Kunze, Juan Madruga, Mike Oaksford, Dan Osherson, David Over, and Marco Ragni, for their help, advice, and criticisms. We are also grateful to Adele Goldberg and the members of her laboratory at Princeton University, who provided feedback on an informal presentation of some of the ideas in this article.

References

Barres, P., & Johnson-Laird, P. N. (2003). On imagining what is true (and what is false). *Thinking & Reasoning, 9,* 1–42. doi:10.1080/13546780244000097

Barrett, R. B., & Stenner, A. J. (1971). The myth of the exclusive "or. *Mind, 80,* 116–121.

Bauer, M. I., & Johnson-Laird, P. N. (1993). How diagrams can improve reasoning. *Psychological Science, 4,* 372–378. doi:10.1111/j.1467-9280.1993.tb00584.x

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

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

Byrne, R. M. J., & Johnson-Laird, P. N. (2009). "If" and the problems of conditional reasoning. *Trends in Cognitive Sciences, 13,* 282–287. doi:10.1016/j.tics.2009.04.003

Chevallier, C., Noveck, I. A., Nazir, T., Bott, L., Lanzetti, V., & Sperber, D. (2008). Making disjunctions exclusive. *Quarterly Journal of Experimental Psychology, 61,* 1741–1760. doi:10.1080/17470210701712960

Chierchia, G., & McConnell-Ginet, S. (1990). *Meaning and grammar*. Cambridge: MIT Press.

Crain, S., & Khlentzos, D. (2007). Is logic innate? *Biolinguistics, 2,* 24–56.

Edghill, E. M. (1928). Categoriae and De Interpretatione. In W. D. Ross (Ed.), *The works of Aristotle*. Oxford: Oxford University Press, Clarendon Press.

Evans, J. St B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences, 7,* 454–459. doi:10.1016/j.tics.2003.08.012

Evans, J. St B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.

Fillenbaum, S. (1974). Or: Some uses. *Journal of Experimental Psychology, 103,* 913–921. doi:10.1037/h0037455

García-Madruga, J. A., Moreno, S., Carriedo, N., Gutiérrez, F., & Johnson-Laird, P. N. (2001). Are conjunctive inferences easier than disjunctive inferences? A comparison of rules and models. *Quarterly Journal of Experimental Psychology, 54A,* 613–632. doi:10.1080/713755974

Geiger, S. M., & Oberauer, K. (2010). Towards a reconciliation of mental model theory and probabilistic theories of conditionals. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 289–307). Oxford: Oxford University Press.

Goldvarg, Y., & Johnson-Laird, P. N. (2000). Illusions in modal reasoning. *Memory & Cognition, 28,* 282–294. doi:10.3758/BF03213806

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Speech acts, Vol. 3, pp. 41–58). New York: Academic Press.

Gualmini, A., Crain, S., & Meroni, L. (2000). Acquisition of disjunction in conditional sentences. In *Proceedings of the Boston University Conference on Language Development* (Vol. 24, pp. 367–378). Medford, MA: Cascadilla Press

Hartshorne, C., Weiss, P., & Burks, A. (Eds.). (1933). *Collected papers of Charles Sanders Peirce: Vol. 4. The simplest mathematics*. Cambridge: Harvard University Press.

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

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.

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

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

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109,* 646–678. doi:10.1037/0033-295X.109.4.646

Johnson-Laird, P. N., Byrne, R. M., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99,* 418–439. doi:10.1037/0033-295X.99.3.418

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review, 111,* 640–661. doi:10.1037/0033-295X.111.3.640

Johnson-Laird, P. N., & Savary, F. (1999). Illusory inferences: A novel class of erroneous deductions. *Cognition, 71,* 191–229. doi:10.1016/S0010-0277(99)00015-3

Juhos, C., Quelhas, C., & Johnson-Laird, P. N. (2012). Temporal and spatial relations in sentential reasoning. *Cognition, 122,* 393–404. doi:10.1016/j.cognition.2011.11.007

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

Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to model-theoretic semantics of natural language, formal logic and discourse representation*. Dordrecht: Kluwer.

Khemlani, S., & Johnson-Laird, P. N. (2009). Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition, 37,* 615–623. doi:10.3758/MC.37.5.615

Khemlani, S., Orenes, I., & Johnson-Laird, P. N. (2012). *Negation: A theory of its meaning, representation, and use*. Manuscript submitted for publication. *Journal of Cognitive Psychology*, In press.

Legrenzi, P., Girotto, V., & Johnson-Laird, P. N. (2003). Models of consistency. *Psychological Science, 14,* 131–137. doi:10.1111/1467-9280.t01-1-01431

Macnamara, J. (1986). *A border dispute: The place of logic in psychology*. Cambridge: MIT Press, Bradford Books.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge: Harvard University Press.

Newstead, S. E., & Griggs, R. A. (1983). The language and thought of disjunction. In J. St B. T. Evans (Ed.), *Thinking and reasoning: Psychological approaches* (pp. 76–106). London: Routledge, Chapman & Hall.

Newstead, S. E., Griggs, R. A., & Chrostowski, J. J. (1984). Reasoning with realistic disjunctives. *Quarterly Journal of Experimental Psychology, 36A,* 611–627. doi:10.1080/14640748408402182

Noveck, I. A., Chierchia, G., Chevaux, F., Guelminger, R., & Sylvestre, E. (2002). Linguistic–pragmatic factors in interpreting disjunctions. *Thinking & Reasoning, 8,* 297–326. doi:10.1080/13546780244000079

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2010). Conditional inference and constraint satisfaction: Reconciling mental models and the probabilistic approach. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 309–333). Oxford: Oxford University Press.

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

Rips, L. J. (1997). Goals for a theory of deduction: Reply to Johnson-Laird. *Minds and Machines, 7,* 409–424.

Roberge, J. (1976a). Effects of negation on adults' disjunctive reasoning abilities. *Journal of General Psychology, 94,* 2328.

Roberge, J. J. (1976b). Reasoning with exclusive disjunction arguments. *Quarterly Journal of Experimental Psychology, 28,* 419–427. doi:10.1080/14640747608400569

Roberge, J. J. (1977). Effects of content on inclusive disjunction reasoning. *Quarterly Journal of Experimental Psychology, 29,* 669–676. doi:10.1080/14640747708400640

Roberge, J. J. (1978). Linguistic and psychometric factors in propositional reasoning. *Quarterly Journal of Experimental Psychology, 30,* 705–716. doi:10.1080/14640747808400695

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah: Erlbaum.

Stenning, K., & van Lambalgen, M. (2008). *Human reasoning and cognitive science*. Cambridge: MIT Press.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005). A dual-process specification of causal conditional reasoning. *Thinking & Reasoning, 11,* 239–278. doi:10.1080/13546780442000178

Walsh, C. R., & Johnson-Laird, P. N. (2004). Co-reference and reasoning. *Memory & Cognition, 32,* 96–106. doi:10.3758/BF03195823

Yang, Y., & Johnson-Laird, P. N. (2000a). How to eliminate illusions in quantified reasoning. *Memory & Cognition, 28,* 1050–1059. doi:10.3758/BF03209353

Yang, Y., & Johnson-Laird, P. N. (2000b). Illusions in quantified reasoning: How to make the impossible seem possible, and vice versa. *Memory & Cognition, 28,* 452–465. doi:10.3758/BF03198560

Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural Language Semantics, 8,* 255–290.