

Reasoning from connectives and relations between entities

Robert Mackiewicz · Philip N. Johnson-Laird

Published online: 15 October 2011

© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract This article reports investigations of inferences that depend both on connectives between clauses, such as *or else*, and on relations between entities, such as *in the same place as*. Participants made more valid inferences from biconditionals—for instance, *Ann is taller than Beth if and only if Beth is taller than Cath*—than from exclusive disjunctions (Exp. 1). They made more valid transitive inferences from a biconditional when a categorical premise affirmed rather than denied one of its clauses, but they made more valid transitive inferences from an exclusive disjunction when a categorical premise denied rather than affirmed one of its clauses (Exp. 2). From exclusive disjunctions, such as *Either Ann is not in the same place as Beth or else Beth is not in the same place as Cath*, individuals tended to infer that all three individuals could be in different places, whereas in fact this was impossible (Exps. 3a and 3b). The theory of mental models predicts all of these results.

Keywords Deductive reasoning · Sentential connectives · Relations · Mental models

Consider the following inferential problem:

Either Ann is taller than Beth or else Beth is taller than Cath, but not both.

So, is it possible that Beth is the tallest of the three?

R. Mackiewicz (✉)
Department of Psychology,
Warsaw School of Social Sciences and Humanities,
Chodakowska 19/31,
03-815, Warsaw, Poland
e-mail: robert.mackiewicz@swps.edu.pl

P. N. Johnson-Laird
Department of Psychology, Princeton University,
Princeton, NJ 08540, USA

The logically correct answer is *yes*, and the validity of the inference depends both on the sentential connective *or else*, which interrelates the two clauses, and on the relations *is taller than* and *is the tallest of*, which interrelate entities. Previous studies have investigated how people understand connectives and make inferences from them, and how people understand and make inferences from relations. But no empirical studies have examined reasoning that hinges both on connectives and on relations, and this novel domain challenges theories of reasoning. The present article addresses this challenge and reports some new phenomena concerning such inferences.

Relations in everyday discourse have various logical properties, of which perhaps the three most important are transitivity, symmetry, and reflexivity (Tarski, 1965, chap. V). A relation such as *is in the same place as* is transitive because if *A is in the same place as B* and *B is in the same place as C*, then *A is in the same place as C*. It is symmetric because if *A is in the same place as B*, then *B is in the same place as A*. And it is reflexive because *A is in the same place as A*. A relation such as *is taller than* is transitive, but it is asymmetric because if *A is taller than B*, then *B is not taller than A*, and it is also irreflexive because *A is not taller than A*. Relations do have other logical properties (see Tarski, 1965), but these properties are recondite and seldom, if ever, relevant to everyday discourse. Indeed, most relational terms in language have no important logical properties—for instance, *A loves B* is neither transitive nor intransitive, neither symmetric nor asymmetric, and neither reflexive nor irreflexive.

Previous psychological studies have investigated how individuals make inferences from transitive relations—for instance, *A is bigger than B* and *B is bigger than C*; therefore, *A is bigger than C* (e.g., Clark, 1969; Huttenlocher, 1968). Transitive inferences occur in everyday life, in intelligence tests, and in inferring economic preferences

(Tversky & Kahneman, 1986). The difficulty of such inferences depends on the number of relations that must be integrated (Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004), the distance between queried elements (Mynatt & Smith, 1977), and the elicitation of extraneous visual images (e.g., Knauff, Fangmeier, Ruff, & Johnson-Laird, 2003). Previous studies have also examined inferences based on two-dimensional spatial relations, temporal relations, and relations between relations, and have shown that the difficulty of inferences also depends on the number of possibilities in which the premises hold (e.g., Byrne & Johnson-Laird, 1989; Carreiras & Santamaria, 1997; Goodwin & Johnson-Laird, 2005, 2006; Schaeken, Johnson-Laird, & d’Ydewalle, 1996). Other studies have shown that some relations are pseudo-transitive—for instance, *A is a blood relative of B*—because individuals tend to make transitive inferences from them unless they are cued to counterexamples (Goodwin & Johnson-Laird, 2008).

The connectives that occur in everyday discourse include *if*, *and*, and *or*. These connectives resemble those of sentential logic (see, e.g., Jeffrey, 1981), but the degree of resemblance is a matter of controversy, especially in the case of *if* (e.g., Evans & Over, 2004; Girotto & Johnson-Laird, 2004; Johnson-Laird & Byrne, 2002; Oaksford & Chater, 2007). Psychologists have carried out many studies of inferences from conditional assertions, *if A then B*, where the values of *A* and *B* are clauses, such as *there is a vowel on a card* (for reviews, see, e.g., Evans & Over, 2004; Johnson-Laird, 2006). These studies, as well as those of other connectives such as *or else*, concern inferences that hinge solely on the connectives, and not on relations between entities that are described in their individual clauses (e.g., Braine & O’Brien, 1998; Johnson-Laird, Byrne, & Schaeken, 1992).

Current accounts of reasoning include those based on formal rules of inference (e.g., Braine & O’Brien, 1998; Rips, 1994), those based on suppositions (e.g., Evans & Over, 2004), those based on probabilistic considerations (e.g., Geiger & Oberauer, 2010; Oaksford & Chater, 2007), and those based on mental models (e.g., Johnson-Laird & Byrne, 1991). None of these accounts has hitherto been applied to inferences that depend on both relations and connectives. Formal rule theories, however, apply to such inferences, because these theories can postulate “meaning postulates” to capture the logical properties of relations. For example, the following meaning postulate captures the transitivity of *is taller than*:

For any x, y, z , if x is taller than y , and y is taller than z , then x is taller than z .

Proponents of such theories have also proposed meaning postulates for two-dimensional spatial reasoning (e.g., Hager, 1984). But the case against meaning postulates

has been presented elsewhere (e.g., Goodwin & Johnson-Laird, 2008). The suppositional theory of Evans and Over (2004) and the probabilistic theory of Oaksford and Chater (2007) have been applied primarily to *if* among the sentential connectives, which these theories analyze in terms of the conditional probability of the *then* clause, given that the *if* clause is true. Although both of these theories might be extended to other connectives and to reasoning that hinges on both relations and connectives, these extensions have yet to be made. Our aim is therefore to follow up the predictions of the theory of mental models rather than to try to pit one theory against the others, because of the uncertainty about their predictions.

The mental model theory—the “model” theory, for short—applies to both relations and connectives (e.g., Johnson-Laird & Byrne, 1991). It is based on three principles: First, a representation of the meaning of an assertion is used to construct mental models of each distinct sort of possibility to which the assertion refers. Second, mental models are iconic insofar as they can be; that is, the structure of a model corresponds to the structure of what it represents (see Peirce, 1958). Third, mental models represent what is true at the expense of what is false (the principle of truth). According to this account, reasoners use the meaning of premises and their knowledge to construct mental models of premises, and depending on whether a conclusion holds in all, most, or some of the models, they draw a conclusion of its necessity, possibility, or probability.

According to the model theory, logical properties such as transitivity are emergent properties from models (e.g., Byrne & Johnson-Laird, 1989; Goodwin & Johnson-Laird, 2005; Huttenlocher, 1968). Hence, the mental model of an assertion such as *Ann is taller than Beth* is an iconic representation of the respective heights of the two individuals, and we use the following sort of diagram to denote such a model:



The further premise, *Beth is taller than Cath*, updates the model:



This model yields the transitive conclusion *Ann is taller than Cath*. No alternative model of the premises refutes this conclusion, so the inference is valid. As Peirce (1958) realized, iconic models have the great advantage that they can yield relations that are not asserted overtly in the premises.

The hypothesis that spatial, temporal, and other sorts of relations are represented in iconic models has received support from behavioral experiments (e.g., Vandierendonck & De Vooght, 1997) and fMRI studies (Goel & Dolan, 2001). But there is evidence that the linguistic form of relational terms, such as *left* or *right*, also influences the conclusions that individuals draw in relational inferences (Van der Henst & Schaeken, 2005). Studies have investigated how individuals construct mental models for spatial relations (Goodwin & Johnson-Laird, 2005; Rauh et al., 2005; Van der Henst, Chevallier, Schaeken, Mercier, & Noveck, 2008). These studies have shown that when descriptions are compatible with more than one spatial layout, individuals tend to construct a particular model; that is, the reasoning system has a preferred model (e.g., Jahn, Knauff, & Johnson-Laird, 2007; Ragni, Knauff, & Nebel, 2005; Rauh et al., 2005), but individuals may also use a single representation that encodes the spatial indeterminacy (e.g., Schaeken, Van der Henst, & Schroyens, 2007; Vandierendonck, Dierckx, & De Vooght, 2004).

The model theory also accounts for sentential reasoning (e.g., Johnson-Laird & Byrne, 1991). For instance, a biconditional of the grammatical form *There is a triangle if and only if there is a circle* has a mental model that represents the presence of the two shapes within the same possibility. Individuals are usually aware that there is another possibility, corresponding to the case in which the *if* clause is false, but they do not normally represent it explicitly (Johnson-Laird & Byrne, 1991, 2002). Instead, they represent this possibility in a model that has no explicit content—that is, an implicit model. The mental models of the preceding biconditional are, accordingly,

$$\Delta \quad \circ$$

...

where the ellipsis denotes the implicit model and each line in the diagram corresponds to a model of a different possibility. When the task is not too burdensome, individuals can flesh out their mental models into *fully explicit* models that represent both possibilities to which the biconditional refers:

$$\Delta \quad \circ$$

$$\text{not-}\Delta \quad \text{not-}\circ$$

where *not-Δ* represents that there is not a triangle—that is, the falsity of the clause.

The principle of truth postulates that mental models do not represent what is false. Hence, the mental models above do not represent the cases that the biconditional rules out as false, i.e., a triangle without a circle, and a circle without a triangle. These cases are constructed by inference from the fully explicit models (Barres & Johnson-Laird, 2003). The principle of truth also applies at a lower level: Mental models do not represent the clauses in the premises that are false in possibilities. For instance, an exclusive disjunction of the grammatical form *Either there is a triangle or else there isn't a circle (but both clauses cannot be true)* yields mental models of the two alternative possibilities:

$$\Delta$$

$$\text{not-}\circ$$

The first of these models does not represent that it is false that *there isn't a circle* in this possibility—that is, *there is a circle*. Likewise, the second model does not represent that it is false that *there is a triangle* in this possibility. Only the fully explicit models represent this information:

$$\Delta \quad \circ$$

$$\text{not-}\Delta \quad \text{not-}\circ$$

Individuals usually do not realize that the disjunction is equivalent to the earlier biconditional, and this phenomenon corroborates the hypothesis that they normally rely on mental models, especially in the case of disjunctions (Johnson-Laird et al., 1992). Experiments have corroborated the model theory of sentential connectives: Inferences are harder with an increase in the number of models that reasoners have to take into account; reasoners tend to overlook models; and reasoners fall into error when the failure to represent what is false matters (see, e.g., Johnson-Laird, 2006, for a review).

The model theory of reasoning based on connectives and relations

We now show how the model theory applies to inferences based on both connectives and relations, and we derive five main predictions. The first prediction concerns the difference between inferences that depend solely on connectives and those that also depend on relations. Consider a biconditional of relations, such as

Ann is in the same place as Beth if and only if Beth is in the same place as Cath.

It calls for the integration of a model of the first clause with a model of the second clause. Reasoners who do not integrate the two models of relations can still infer that Beth is in the same place as Cath from the categorical assertion that Ann is the same place as Beth. Such inferences

interrelate two individuals according to a single clause in a compound premise, and we refer to them as yielding a “nontransitive” conclusion. In order to make a transitive inference that Ann is in the same place as Cath, reasoners have to integrate the two models into a single model in which all three individuals are in the same place. The first prediction is accordingly that reasoners should make nontransitive inferences more often than transitive inferences.

The second prediction concerns integrated mental models and sentential connectives. The integrated mental models of the biconditional above are as follows, where the rectangle denotes a single place and the ellipsis denotes an implicit model representing the possibility in which both clauses are false:

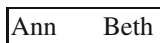


. . .

In contrast, the logically equivalent exclusive disjunction with a negative second clause

Either Ann is in the same place as Beth or else Beth is not in the same place as Cath.

has the following mental models:



These two mental models cannot be integrated because they represent two alternative possibilities. The processing capacity of working memory is limited (e.g., Capon, Handley, & Dennis, 2003), and a single mental model with content imposes less of a load on working memory than do two such mental models (Barrouillet & Lecas, 1999). The second prediction of the theory is therefore that reasoning, in general, should be easier to make from biconditionals than from exclusive disjunctions.

The third prediction is that inferences should be easier to make from symmetric relations, such as *is in the same place as*, than from asymmetric relations, such as *is taller than*, because the former do not require participants to remember which entity is the subject of a relation and which entity is its object, whereas the latter calls for this information to be kept in mind.

The fourth prediction concerns inferences based on a compound premise containing a connective and a

categorical premise, such as *Ann is taller than Beth*. The prediction concerns an interaction, and we describe it first for biconditionals and then for disjunctions. The mental models of a biconditional represent the possibility in which both clauses are true, and implicitly the possibility in which both clauses are false. However, previous research has shown that reasoners are quite likely to flesh out the implicit model of a biconditional explicitly (see, e.g., Johnson-Laird et al., 1992). Hence, there is not a marked difference between drawing a conclusion when a clause is affirmed and drawing a conclusion when a clause is denied. It follows that these two sorts of nontransitive inference should be of comparable difficulty:

Ann is taller than Beth if and only if Beth is taller than Cath.
Ann is taller than Beth.
What follows? [Beth is taller than Cath.]

and

Ann is taller than Beth if and only if Beth is taller than Cath.
Ann is not taller than Beth.
What follows? [Beth is not taller than Cath.]

However, if individuals form an integrated mental model of both clauses in the biconditional, which represents the relative heights of all three individuals, they can draw the transitive conclusion that Ann is taller than Cath. The cost of this integration is that they will be much less likely to flesh out the implicit model of the case in which both clauses are false. Hence, it follows that the transitive inference from the premises in the first example above should be easier to make than the transitive inference from the premises in the second example, which yields the conclusion that Ann is not taller than Cath.

The two disjunctive inferences logically equivalent to the examples above are as follows:

Ann is taller than Beth or else Beth is not taller than Cath.
Ann is taller than Beth.
What follows?

and

Ann is taller than Beth or else Beth is not taller than Cath.
Ann is not taller than Beth.
What follows?

The disjunction has two mental models of alternative possibilities. In the first inference, the categorical premise matches the first model, and a conclusion can be drawn

only if this model is fleshed out to represent the falsity of the second clause. But in the second inference, the categorical premise eliminates the first model of the disjunction, and the second model yields an immediate nontransitive conclusion. Both nontransitive and transitive inferences should therefore be easier to make when a premise denies a clause of a disjunction than when a premise affirms a clause of a disjunction.

The predicted patterns of difficulty in these inferences is complicated, so we summarize the interaction:

- When a premise affirms a clause in a biconditional, transitive inferences should be easier to make than when a premise denies a clause in a biconditional. (For the reason that we explained earlier, this prediction does not apply to nontransitive inferences.)
- When a premise denies a clause in a disjunction, nontransitive and transitive inferences should be easier to make than when a premise affirms a clause in a disjunction.

The fifth prediction is that systematic fallacies should occur when inferences about possibilities depend on overcoming the principle of truth and representing what is false. Previous studies have shown that such fallacies occur in various domains and can be sufficiently compelling to yield “illusory inferences” (see, e.g., Johnson-Laird, Legrenzi, Girotto, & Legrenzi, 2000; Johnson-Laird & Savary, 1999; Khemlani & Johnson-Laird, 2009). As an example, consider this problem based on an exclusive disjunction:

Either Adam is in the same place as Beth or else Beth is not in the same place as Cath.

Is it possible that Adam and Beth are alone in the same place?

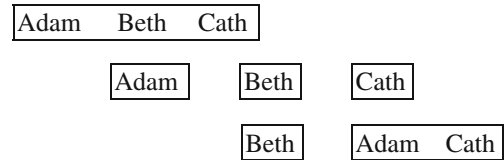
The mental models of the premise are as follows:

Adam Beth

Beth Cath

The first of these models predicts that individuals should respond “Yes.” The fully explicit models of the premise take into account that one clause is true and that the other clause is false. Hence, when Adam is in the same place as Beth, it is false that Beth is not in the same place as Cath—that is, Beth *is* in the same place as Cath—so all three of them are in the same place. When Beth is not in the same place as Cath, then it is false that Adam is in the same place as Beth, and so either all three

are in different places, or Adam and Cath could be in the same place. The fully explicit models of the disjunction are therefore as follows:



Each line in the diagram above represents a different possibility. Hence, the valid inference is “No, Adam and Beth cannot be alone in the same place,” and the previous inference is fallacious. Now, consider the case in which the premise is followed by the question

Is it possible that Beth and Cath are alone in the same place?

The mental models yield the answer “No.” And, in this case, the fully explicit models also yield this answer. The inference is a control, which the model theory predicts should elicit the correct response. Analogous predictions apply to inferences based on biconditionals.

Our experiments were designed to test the five predictions that we have described. They used the two simple binary relations that we have illustrated above: *is in the same place as* and *is taller than*. In a preliminary study, participants had to draw diagrams to represent these relations in affirmative and negative assertions. In almost every case, they used vertical bars or schematic drawings of humans to represent heights, and they used boxes or parentheses to represent different places. The experiments were carried out in Polish, which has unambiguous ways to express an exclusive disjunction, *albo A albo B* (*either A or else B*), and a biconditional, *A wtedy i tylko wtedy, gdy B* (*A if and only if B*), where *A* and *B* are clauses. Both of these expressions are probably used more frequently in daily life in Polish than in English (see the corpus of the Polish language: korpus.pwn.pl). For the rest of this article, we use the term “disjunction” to refer to an exclusive disjunction.

Experiment 1

The aim of this experiment was to test the theory’s predictions about inferences such as

Ann is taller than Beth if and only if Beth is taller than Cath.

Ann is taller than Beth.

These premises validly yield a simple “nontransitive” conclusion—that is, one concerning only a single clause in the biconditional:

Beth is taller than Cath.

But individuals who integrate their mental model of the biconditional can also draw the transitive conclusion

Ann is taller than Cath.

The theory predicts that both sorts of inference should be easier to make from biconditional premises than from logically equivalent disjunctive premises with a negative second clause. The experiment also tested the prediction that reasoning from a symmetric relation (*is in the same place as*) should be easier than reasoning from an asymmetric relation (*is taller than*). The experiment used a task in which the participants drew their own conclusions from premises.

Method

Participants A group of 32 students of the Warsaw School of Social Sciences and Humanities participated in the experiment as a course requirement. None of them had any training in logic beyond an introductory course, which is obligatory in all Polish universities.

Design and materials The participants acted as their own controls and carried out two inferences of each of eight sorts. One inference of each sort concerned a symmetric relation and the other inference concerned an asymmetric relation. Every inference concerned three referents, in the order *A–B* and *B–C*, in the two clauses of the compound premise. In half of the trials, both clauses were affirmative, and in the other half, the first clause was affirmative and the second clause was negative. The resulting pairs of relations were presented either in a biconditional or a disjunction. Each sentence was accompanied by a categorical premise that affirmed either the first or the second clause in the compound premise. The resulting 16 inferences all yielded a valid conclusion, so the experiment also contained two filler inferences that did not yield a valid conclusion. All of the experimental inferences in [Experiment 1](#) are presented in [Appendix A](#). All of the inferences were about different people, and three different names were assigned at random to them.

Procedure The participants were tested in four small groups, and the problems were presented in booklets. The instructions stated that the aim of the study was to see how people understand assertions describing different sorts of relations. The key instructions were as follows:

The experiment concerns a number of situations, which are each described in two sentences. Your task is to draw a

correct conclusion to each inference, that is, a conclusion that must be true given that the two sentences . . . are true. Please analyze carefully each situation, and write down in the appropriate place the conclusion that you think is correct. If you are convinced that no correct conclusion can be drawn in a particular case, write this as your answer.

The participants wrote down their answers in the space provided below each pair of premises.

Results and discussion

The overall percentages of valid conclusions, both nontransitive and transitive, were as follows:

Biconditionals of symmetric relations:	84%
Biconditionals of asymmetric relations:	72%
Disjunctions of symmetric relations:	55%
Disjunctions of asymmetric relations:	46%

The percentages of valid conclusions to each problem are shown in [Appendix A](#), and the preceding pattern did not differ reliably whether the categorical premise affirmed the first or the second clause in the compound premise. As the model theory predicts, the participants drew more valid conclusions from biconditionals (78%) than from disjunctions (51%; Wilcoxon test, $z = 4.44$, $p < .001$). Likewise, as the theory predicts, they drew more valid conclusions from symmetric relations (70%) than from asymmetric relations (59%; Wilcoxon test, $z = 2.38$, $p = .009$). There was no reliable interaction between the two sorts of connective and the two sorts of relation (Wilcoxon test, $z = 0.05$, $p = .96$, two tailed). The *transitive* inferences could be drawn from four asymmetric and seven symmetric problems (see [Appx. A](#)). Only 32% of the answers to those inferences were correct transitive conclusions, and only 5 out of 32 participants did not give any transitive conclusion in the task. These 5 individuals may not have built any integrated models of the premises. To compare transitive inferences from biconditionals and disjunctions, we examined only those pairs with the same categorical premise and the same conclusion. As the model theory predicts, the participants drew more valid transitive conclusions from biconditionals (41%) than from disjunctions (11%; Wilcoxon test, $z = 3.05$, $p = .001$). But there was no reliable difference in valid transitive conclusions from symmetric (28%) and asymmetric (23%; Wilcoxon test, $z = .41$, $p = .34$) relations. The interaction between the two variables was not significant (Wilcoxon test, $z = 1.33$, $p = .18$, two tailed). Overall, the results corroborated the model theory: The participants made a greater number of valid inferences about relations from biconditionals than from disjunctions, both for nontransitive conclusions from a single clause and for transitive inferences that combined individuals in the separate clauses.

Experiment 2

The main aim of this experiment was to test the model theory's predicted interaction between the categorical premise's role and the connective in the compound premise: Affirmations should yield more transitive inferences from biconditional premises, whereas denials should yield more transitive inferences from disjunctive premises. With that aim, the experiment used only the more difficult, asymmetric relation *is taller than*, and each biconditional occurred with equivalent inferences from disjunctions. The numbers of negations were matched in the biconditional and disjunctive inferences.

Method

Participants A group of 31 psychology undergraduates from the same population as before took part in the study as a course requirement. Two of the participants failed to give conclusions to all the inferences, and two others misunderstood the task, so the analysis of the results concerned the remaining 27 participants.

Design, materials, and procedure Each participant made inferences from a biconditional and from a disjunction of two relations, both of which concerned three referents in the order *A–B* and *B–C* in the two clauses of the compound premise. The biconditionals consisted of either two affirmative or two negative relations:

A is taller than B if and only if B is taller than C

and

A is not taller than B if and only if B is not taller than C.

The disjunctions had the same fully explicit models as these two biconditionals:

A is taller than B or else B is not taller than C

and

A is not taller than B or else B is taller than C.

The categorical premise either affirmed or denied one clause in the compound premise, either the first or the second clause. These manipulations resulted in eight biconditional and eight disjunctive inferences, half of them with an affirmative categorical and half with a negative categorical. Additional two filler inferences did not yield a valid conclusion. All of the resulting 18 inferences referred to three different individuals. Four different orders of inferences were assigned at random to the participants in the study. The full set of experimental inferences is stated in [Appendix B](#). The procedure and instructions were identical to those in the previous experiment, and the participants had to draw their own conclusions or respond that “nothing follows.”

Results

Table 1 presents the percentages of valid nontransitive and transitive conclusions for both sorts of compound premise, and the results for the individual inferences are presented in [Appendix B](#). Overall, the participants drew more nontransitive inferences (46%) than transitive inferences (11%; Wilcoxon test, $z = 3.77$, $p < .001$). They drew more valid inferences of both sorts from biconditionals (78%) than from disjunctions (36% valid; Wilcoxon test, $z = 4.24$, $p < .001$). The frequencies of transitive conclusions did not differ overall, either between biconditional and disjunctive inferences (Wilcoxon test, $z = 0.96$, $p = .33$, two tailed) or between inferences based on categorical premises affirming or negating clauses in the compound premise (Wilcoxon test, $z = 1.54$, $p = .12$, two tailed). The predicted interaction between these two variables was reliable for transitive inferences overall (Wilcoxon test, $z = 3.07$, $p = .001$), but although the trend for nontransitive inferences from disjunctions was in the predicted direction, it was not reliable (Wilcoxon test, $z = 1.18$, $p = .24$). Hence, with this one exception, the results corroborated the model theory's predictions for these inferences.

Table 1 The percentages of valid inferences from biconditionals and disjunctions paired with categorical premises that affirm or deny one of their clauses

	Biconditional Premise		Disjunctive Premise	
	Categorical premise affirms a clause in biconditional	Categorical premise denies a clause in biconditional	Categorical premise affirms a clause in disjunction	Categorical premise denies a clause in disjunction
Valid nontransitive conclusions	63	68	22	32
Valid transitive conclusions	22	4	6	12

Nontransitive conclusions assert one of the clauses in the compound premise, whereas transitive conclusions assert a relation between items in both clauses

Experiments 3a and 3b

Our final two experiments tested the model theory's prediction of systematic fallacies when inferences about possibilities depend on overcoming the principle of truth and representing what is false. The experiments examined the predicted fallacies and control inferences, and both sorts of inference depended on connectives and relations. Experiment 3a was carried out online, and Experiment 3b was a face-to-face test. Otherwise, the two experiments were identical, so we report them together.

Method

Participants We carried out Experiment 3a over the Internet, and 42 participants were recruited from among first-year psychology students enrolled in a course of economic psychology in weekend classes at the Warsaw School of Social Sciences and Humanities. Their mean age was 39 years, and 69% had a higher-education diploma. We carried out Experiment 3b in a direct test of 40 undergraduate psychology students from the same population as in our earlier experiments.

Design and materials The participants acted as their own controls, and they made four inferences about possibilities from a biconditional and four inferences about possibilities from a disjunction. The premises concerned three referents in the order $A-B$ and $B-C$ in the two clauses. There were two different biconditional premises and two different disjunctive premises (see Table 2 below). Each compound premise was paired on separate trials with two different questions about possible relations. One question was "illusory": Mental models yielded a fallacious answer, and the valid answer depended on fully explicit models. The other question was a control: Mental models and fully explicit models both yielded the same valid answer. In four problems, the valid answer was, "yes", and in the other four the valid answer was, "no". Each inference used three different names of individuals, assigned once at random to the premises. In Experiment 3a, a single set of problems was presented in a different random order to each of the participants. In Experiment 3b, three different sets of eight triples of common names were assigned at random to the problems. The three sets of problems were allocated at random to the participants and were presented to each of them in a random order.

Procedure In Experiment 3a, the participants were recruited by e-mail, and a commercial website (www.wbadanie.pl) was the server for the experiment. This site placed the usual constraints on Web experiments—for instance, the test could be taken only once from a computer. In Experiment 3b, the participants were tested in two groups but worked individually. In both experiments, the key instructions were "On each

trial, you will see a sentence describing a situation, and a question. Your task is to answer the question by deciding whether or not the situation it describes is possible given the truth of the sentence." Online, the participants chose their answers by selecting one of two buttons: "Yes, it is possible" or "No, it is impossible." Each problem was presented on a single screen and had to be answered in order to go on to the next inference. In the face-to-face experiment, the participants ticked one of two boxes under the problem, corresponding to the two options.

Results and discussion

Table 2 presents the percentages of "yes" and "no" answers to each inference in the two experiments. The results were similar in both studies. As the model theory predicts, the participants were correct more often for control inferences (74% and 74% for the two experiments, respectively) than for the illusory inferences (45% and 44% for the two experiments). The difference was reliable in both experiments (Wilcoxon tests, $z = 3.99$ and 4.32 , $p < .001$, for both tests). The participants were more accurate with biconditional inferences (66% and 68% for the two experiments) than with disjunctive inferences (53% and 51%; Wilcoxon tests, $z = 2.24$ and 2.29 , $p = .012$, for both tests). Correct answers were more frequent for biconditional illusory inferences (54% in Exp. 3a and 59% in Exp. 3b) than for disjunctive illusory inferences (36% in Exp. 3a and 30% in Exp. 3b; Wilcoxon tests, $z = 2.17$, $p = .03$, and $z = 3.29$, $p = .001$). This performance with the disjunctive illusory inferences was reliably worse than the chance probability of 50% (Wilcoxon tests, $z = 2.85$, $p < .003$, and $z = 1.85$, $p < .04$, for Exps. 3a and 3b, respectively).

Both experiments corroborated the model theory's prediction. When mental models yield a correct response, inferences are reliably easier to make than when only fully explicit models yield a correct response. The results also confirmed, once again, that inferences tend to be easier to make from biconditionals than from disjunctions. The similarity between the results of the online and face-to-face experiments is reassuring about the use of Web-based studies of reasoning.

General discussion

The present studies examined a novel sort of reasoning that depends both on sentential connectives such as *or else*, which interrelate clauses, and on relations such as *is in the same place as*, which interrelate entities. Connectives and relations have hitherto been studied separately, so the aim was to test whether a theory based on the idea that individuals construct mental models of possibilities could

Table 2 The problems in Experiments 3a (online) and 3b (face-to-face), their mental models and fully explicit models with square brackets demarcating separate places, and the percentages of valid (in bold) and invalid responses

Compound Premises With Illusory and Control Questions					Mental Models and Fully Explicit Models for the Premises
Inferences With Disjunctive Compound Premises					
1. A is not in the same place as B or else B is not in the same place as C.					Mental models
					[A] [B]
Illusory question:					[B] [C]
Is it possible that all three are in different places?					Fully explicit models
3a:	Yes	60	No	40	[A] [B C]
3b:	Yes	63	No	37	[A B] [C]
Control question:					
Is it possible that A and C are in different places?					
3a:	Yes	88	No	12	
3b:	Yes	88	No	13	
2. A is in the same place as B or else B is not in the same place as C.					Mental models
					[A B]
Illusory question:					[B] [C]
Is it possible that A and B alone are in the same place?					Fully explicit models
3a:	Yes	69	No	31	[A B C]
3b:	Yes	78	No	23	[A] [B] [C]
Control question:					[B] [A C]
Is it possible that B and C alone are in the same place?					
3a:	Yes	48	No	52	
3b:	Yes	45	No	55	
Inferences With Biconditional Compound Premises					
3. A is in the same place as B if and only if B is in the same place as C.					Mental models
					[A B C]
Illusory question:					...

Table 2 (continued)

Compound Premises With Illusory and Control Questions	Mental Models and Fully Explicit Models for the Premises
Is it possible that A and C are the only two in the same place?	Fully explicit models
3a: Yes 52 No 48	[A B C]
3b: Yes 45 No 55	[A] [B] [C]
Control question:	[B] [A C]
Is it possible that A and B are the only two in the same place?	
3a: Yes 17 No 83	
3b: Yes 30 No 70	
4. A is in the same place as B if and only if B is not in the same place as C.	Mental models
	[A B] [C]
Illusory question:	...
Is it possible that B and C are in the same place?	Fully explicit models
3a: Yes 55 No 45	[A B] [C]
3b: Yes 73 No 29	[A] [B C]
Control question:	
Is it possible that A is not in the same place as C	
3a: Yes 74 No 26	
3b: Yes 85 No 15	

explain why some of these inferences are harder to make than others. The model theory predicts a difference between inferences from biconditionals, such as

Ann is in the same place as Beth if and only if Beth is in the same place as Cath.

and inferences based on logically equivalent exclusive disjunctions,

Ann is in the same place as Beth or else Beth is not in the same place as Cath.

The biconditional has a single mental model of a possibility and an implicit model with no content, whereas the disjunction has two mental models for each alternative possibility. It follows that inferences in general should be easier to make from biconditionals than from disjunctions. Experiments 1 and 2 corroborated this prediction.

Reasoners do not have to keep track of the arguments of symmetrical relations, because *A is in the same place as B* is equivalent to *B is in the same place as A*. But they do have to keep track of the arguments of asymmetrical relations, because *A is taller than B* is not equivalent to *B is taller than A*. Hence, inferences from symmetrical relations should be easier to make than inferences from asymmetrical relations. **Experiment 1** also corroborated this prediction.

Consider the following inferential problem:

Ann is taller than Beth if and only if Beth is taller than Cath.
Ann is taller than Beth.
What follows?

Individuals could build a mental model that represents each clause separately, or they could integrate the information from the two clauses. In the first case, they could draw the nontransitive conclusion *Beth is taller than Cath*. In the

second case, they can draw the transitive conclusion *Ann is taller than Cath*. The integration of information from separate clauses takes work, so the theory predicts that reasoners should be more likely to infer nontransitive than transitive conclusions. **Experiment 2** corroborated this prediction.

The theory also predicts an interaction for transitive inferences. Transitive inferences should be easier to make from biconditionals when a premise affirms one of their clauses, but they should be easier to make from disjunctions when a premise denies one of their clauses. **Experiment 2** corroborated this interaction in a study of inferences from the relation *is taller than*. The principle underlying the prediction about transitive inferences was straightforward—an affirmation of a clause in a biconditional yields the transitive conclusion from an integrated mental model, whereas a denial of a clause does not—it calls for models to be fleshed out in a fully explicit way. Likewise, a denial of a clause in disjunction eliminates one of its models, and the conclusion follows from an integrated version of its other model, whereas an affirmation of a clause calls for its model to be fleshed out with a fully explicit model of the falsity of the second clause. The complete motivation for this prediction is described in the introduction.

The principle of truth, which is embodied in mental models, predicts that it is harder to make valid inferences that depend on taking into account what is false. For example, reasoners should tend to respond “yes” to the following problem:

Ann is in the same place as Beth or else Beth is not in the same place as Cath.

Is it possible that Ann and Beth alone are in the same place?

The mental models of the premise yield an affirmative answer, because they represent two possibilities: Ann and Beth in the same place, and Beth and Cath in different places. The fully explicit models show that the affirmative answer is wrong, because when Ann and Beth are in the same place, so too is Cath (from the falsity of the second clause in the disjunction). Individuals have difficulties with such inferences, but, as the theory also predicts, they do reliably better with control inferences for which the valid answer is not affected by a failure to represent what is false (Exps. 3a and 3b).

In sum, five main phenomena corroborated the predictions of the model theory:

- Individuals made more valid nontransitive than transitive inferences.
- They made more valid inferences from biconditionals than from disjunctions.
- They made valid inferences more often from symmetric relations, such as *is in the same place as*, than from asymmetric relations, such as *is taller than*.
- They made transitive inferences more often when a categorical premise affirmed one of the clauses in a

biconditional than when it denied one of the clauses, but the difference switched round with disjunctions.

- They made more valid inferences from mental models than from fully explicit models that represented both what is true and what is false.

The results accordingly support the theory that reasoning from relations and connectives depends on mental models, which embody the principle of truth. An alternative theory could in principle account for these results. It is always possible to devise such an account after the fact, but to the best of our knowledge, no other current theory of reasoning makes all five predictions.

Author note This research was supported by National Science Centre grant 3636/B/H03/2011/40 to the first author for studying ambiguities in reasoning, and by National Science Foundation Grant SES 0844851 to the second author for studying deductive and probabilistic reasoning. We thank Sangeet Khemlani and Max Lotstein for advice.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix A

Inferences in **Experiment 1** and the percentages (in parentheses) of participants who spontaneously drew valid transitive conclusions (possible only for four asymmetric and seven symmetric problems) and valid nontransitive conclusions about single clauses

Biconditional Inferences	Disjunctive Inferences
Asymmetric Relations	
1	1'
A is taller than B if and only if B is taller than C.	A is taller than B or else B is taller than C.
A is taller than B.	A is taller than B.
Valid transitive conclusion: A is taller than C. (34)	Valid nontransitive conclusion: B is not taller than C. (53)
Valid nontransitive conclusion: B is taller than C. (50)	
2	2'
A is taller than B if and only if B is taller than C.	A is taller than B or else B is taller than C.
B is taller than C.	B is taller than C.
Valid transitive conclusion: A is taller than C. (16)	Valid nontransitive conclusion: A is not taller than B. (41)
Valid nontransitive conclusion: A is taller than B. (75)	

3
 A is taller than B if and only if B is not taller than C.
 A is taller than B.
 Valid nontransitive conclusion:
 B is not taller than C. (59)

4
 A is taller than B if and only if B is not taller than C.
 B is not taller than C.
 Valid nontransitive conclusion:
 A is taller than B. (53)

Symmetric Relations

5
 A is in the same place as B if and only if B is in the same place as C.
 A is in the same place as B.
 Valid transitive conclusion:
 A is in the same place as C. (47)
 Valid nontransitive conclusion:
 B is in the same place as C. (38)

6
 A is in the same place as B if and only if
 B is in the same place as C.
 B is in the same place as C.
 Valid transitive conclusion:
 A is in the same place as C. (53)
 Valid nontransitive conclusion:
 A is in the same place as B. (38)

7
 A is in the same place as B if and only if
 B is not in the same place as C.
 A is in the same place as B.
 Valid transitive conclusion:
 A is not in the same place as C. (31)
 Valid nontransitive conclusion:
 B is not in the same place as C. (47)

8
 A is in the same place as B if and only if
 B is not in the same place as C.
 B is not in the same place as C.

3'
 A is taller than B or else B is not taller than C.
 A is taller than B.
 Valid transitive conclusion:
 A is taller than C. (12)
 Valid nontransitive conclusion:
 B is taller than C. (19)

4'
 A is taller than B or else B is not taller than C.
 B is not taller than C.
 Valid transitive conclusion:
 A is not taller than C. (19)
 Valid nontransitive conclusion:
 A is not taller than B. (41)

5'
 A is in the same place as B or else B or else
 B is in the same place as C.
 A is in the same place as B.
 Valid transitive conclusion:
 A is not in the same place as C. (16)
 Valid nontransitive conclusion:
 B is not in the same place as C. (59)

6'
 A is in the same place as B or else
 B is in the same place as C.
 B is in the same place as C.
 Valid transitive conclusion:
 A is not in the same place as C. (31)
 Valid nontransitive conclusion:
 A is not in the same place as B. (44)

7'
 A is in the same place as B or else
 B is not in the same place as C.
 A is in the same place as B.
 Valid transitive conclusion:
 A is in the same place as C. (9)
 Valid nontransitive conclusion:
 B is in the same place as C. (25)

8'
 A is in the same place as B or else
 B is not in the same place as C.
 B is not in the same place as C.

Valid transitive conclusion:
 A is not in the same place as C. (16)
 Valid nontransitive conclusion:
 A is in the same place as B. (66)

Valid nontransitive conclusion:
 A is not in the same place as B. (38)

Appendix B:

Inferences in [Experiment 2](#) and the percentages (in parentheses) of participants who spontaneously drew valid transitive conclusions and valid nontransitive conclusions

Biconditional Inferences	Disjunctive Logical Equivalents
1 A is taller than B if and only if B is taller than C. A is taller than B. Valid transitive conclusion: A is taller than C. (33) Valid nontransitive conclusion: B is taller than C. (59)	1' A is taller than B or else B is not taller than C. A is taller than B. Valid transitive conclusion: A is taller than C. (11) Valid nontransitive conclusion: B is taller than C. (26)
2 A is taller than B if and only if B is taller than C. B is taller than C. Valid transitive conclusion: A is taller than C. (22) Valid nontransitive conclusion: A is taller than B. (70)	2' A is taller than B or else B is not taller than C. B is taller than C. Valid transitive conclusion: A is taller than C. (18) Valid nontransitive conclusion: A is taller than B. (18)
3 A is taller than B if and only if B is taller than C. A is not taller than B. Valid transitive conclusion: A is not taller than C. (4) Valid nontransitive conclusion: B is not taller than C. (67)	3' A is taller than B or else B is not taller than C. A is not taller than B. Valid transitive conclusion: A is not taller than C. (4) Valid nontransitive conclusion: B is not taller than C. (48)
4 A is taller than B if and only if B is taller than C. B is not taller than C. Valid transitive conclusion: A is not taller than C. (11) Valid nontransitive conclusion: A is not taller than B. (67)	4' A is taller than B or else B is not taller than C. B is not taller than C. Valid transitive conclusion: A is not taller than C. (0) Valid nontransitive conclusion: A is not taller than B. (11)
5 A is not taller than B if and only if B is not taller than C.	5' A is not taller than B or else B is taller than C.

A is not taller than B.	A is not taller than B.
Valid transitive conclusion:	Valid transitive conclusion:
A is not taller than C. (7)	A is not taller than C. (7)
Valid nontransitive conclusion:	Valid nontransitive conclusion:
B is not taller than C. (52)	B is not taller than C. (22)
6	6'
A is not taller than B if and only if B is not taller than C.	A is not taller than B or else B is taller than C.
B is not taller than C.	B is not taller than C.
Valid transitive conclusion:	Valid transitive conclusion:
A is not taller than C. (26)	A is not taller than C. (15)
Valid nontransitive conclusion:	Valid nontransitive conclusion:
A is not taller than B. (70)	A is not taller than B. (26)
7	7'
A is not taller than B if and only if B is not taller than C.	A is not taller than B or else B is taller than C.
A is taller than B.	A is taller than B.
Valid transitive conclusion:	Valid transitive conclusion:
A is taller than C. (0)	A is taller than C. (11)
Valid nontransitive conclusion:	Valid nontransitive conclusion:
B is taller than C. (70)	B is taller than C. (33)
8	8'
A is not taller than B if and only if B is not taller than C.	A is not taller than B or else B is taller than C.
B is taller than C.	B is taller than C.
Valid transitive conclusion:	Valid transitive conclusion:
A is taller than C. (0)	A is taller than C. (7)
Valid nontransitive conclusion:	Valid nontransitive conclusion:
A is taller than B. (67)	A is taller than B. (30)

References

- Barres, P. E., & Johnson-Laird, P. N. (2003). On imaging what is true (and what is false). *Thinking & Reasoning*, *9*, 1–42.
- Barrouillet, P., & Lecas, J.-F. (1999). Mental models in conditional reasoning and working memory. *Thinking & Reasoning*, *5*, 289–302.
- Braine, M. D. S., & O'Brien, D. P. (Eds.). (1998). *Mental logic*. Mahwah: Erlbaum.
- Byrne, R. M. J., & Johnson-Laird, P. N. (1989). Spatial reasoning. *Journal of Memory and Language*, *28*, 564–575.
- Capon, A., Handley, S., & Dennis, I. (2003). Working memory and reasoning: An individual differences perspective. *Thinking & Reasoning*, *9*, 203–244.
- Carreiras, M., & Santamaría, C. (1997). Reasoning about relations: Spatial and nonspatial problems. *Thinking & Reasoning*, *3*, 191–208.
- Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review*, *76*, 387–404.
- Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
- Geiger, S., & 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.
- Giroto, V., & Johnson-Laird, P. N. (2004). The probability of conditionals. *Psychologia*, *47*, 207–225.
- Goel, V., & Dolan, R. J. (2001). Functional neuroanatomy of three term relational reasoning. *Neuropsychologia*, *39*, 901–909.
- Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review*, *112*, 468–493. doi:10.1037/0033-295X.112.2.468
- Goodwin, G. P., & Johnson-Laird, P. N. (2006). Reasoning about the relations between relations. *Quarterly Journal of Experimental Psychology*, *59*, 1047–1069. doi:10.1080/02724980543000169
- Goodwin, G. P., & Johnson-Laird, P. N. (2008). Transitive and pseudo-transitive inferences. *Cognition*, *108*, 320–352. doi:10.1016/j.cognition.2008.02.010
- Hagert, G. (1984). Modeling mental models: Experiments in cognitive modeling spatial relations. In T. O'Shea (Ed.), *Advances in artificial intelligence* (pp. 179–188). Amsterdam: North-Holland.
- Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review*, *75*, 550–560.
- Jahn, G., Knauff, M., & Johnson-Laird, P. N. (2007). Preferred mental models in reasoning about spatial relations. *Memory & Cognition*, *35*, 2075–2087.
- Jeffrey, R. (1981). *Formal logic: Its scope and limits*. New York: McGraw-Hill.
- 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.
- Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, *99*, 418–439.
- Johnson-Laird, P. N., Legrenzi, P., Giroto, V., & Legrenzi, M. (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.
- Khemlani, S., & Johnson-Laird, P. N. (2009). Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, *37*, 615–623.
- Knauff, M., Fangmeier, T., Ruff, C. C., & Johnson-Laird, P. N. (2003). Reasoning, models, and images: Behavioral measures and cortical activity. *Journal of Cognitive Neuroscience*, *4*, 559–573.
- Mynatt, B. T., & Smith, K. H. (1977). Constructive processes in linear order problems revealed by sentence study times. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 353–374.
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality*. Oxford: Oxford University Press.
- Peirce, C. S. (1958). In C. Hartshorne, P. Weiss, & A. Burks (Eds.), *Collected papers of Charles Sanders Peirce*. Cambridge: Harvard University Press.
- Ragni, M., Knauff, M., & Nebel, B. (2005). A computational model for spatial reasoning with mental models. In B. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (pp. 1064–1070). Mahwah: Erlbaum.
- Rauh, R., Hagen, C., Knauff, M., Kuss, T., Schlieder, C., & Strube, G. (2005). Preferred and alternative mental models in spatial reasoning. *Spatial Cognition and Computation*, *5*, 239–269. doi:10.1080/13875868.2005.9683805
- Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge: MIT Press.
- Schaeken, W., Johnson-Laird, P. N., & d'Ydewalle, G. (1996). Mental models and temporal reasoning. *Cognition*, *60*, 205–234.

- Schaeken, W. S., Van der Henst, J.-B., & Schroyens, W. (2007). The mental models theory of relational reasoning: Premises' relevance, conclusions' phrasing, and cognitive economy. In W. Schaeken, A. Vandierendonck, W. Schroyens, & G. d'Ydewalle (Eds.), *The mental models theory of reasoning: Refinements and extensions* (pp. 129–149). Mahwah: Erlbaum.
- Tarski, A. (1965). *Introduction to logic and to the methodology of deductive sciences* (3rd ed.). New York: Oxford University Press.
- Tversky, A., & Kahneman, D. (1986). Rational choice and framing of decisions. *Journal of Business*, *59*, 251–278.
- Van der Henst, J.-B., Chevallier, C., Schaeken, W., Mercier, H., & Noveck, I. (2008). Dealing with indeterminacy in spatial descriptions. *Psychological Research*, *72*, 553–566.
- Van der Henst, J.-B., & Schaeken, W. (2005). The wording of conclusions in relational reasoning. *Cognition*, *97*, 1–22.
- Vandierendonck, A., & De Vooght, G. (1997). Working memory constraints on linear reasoning with spatial and temporal contents. *Quarterly Journal of Experimental Psychology*, *50*, 803–820.
- Vandierendonck, A., Dierckx, V., & De Vooght, G. (2004). Mental model construction in linear reasoning: Evidence for the construction of initial annotated models. *Quarterly Journal of Experimental Psychology*, *57A*, 1369–1391.
- Viskontas, I. V., Morrison, R. G., Holyoak, K. J., Hummel, J. E., & Knowlton, B. J. (2004). Relation integration, inhibition, and analogical reasoning in older adults. *Psychological Review*, *19*, 581–591.