Psicológica (2009), 30, 217-243.

Reasoning with 'Unless' counterfactual conditionals

Juan A. García-Madruga[*1], Sergio Moreno-Ríos[2], Ana Cristina Quelhas[3] & Csongor Juhos[3]

[1]*UNED, Madrid, Spain*; [2]*Granada University, Spain;*
[3] *ISPA, Lisbon, Portugal*

This article tackles factual and counterfactual 'unless' expressions such as 'Virginia will not pass the exam unless she works harder' and 'Virginia would not have passed the exam unless she had worked harder'. 'Unless' is a negative conditional that is semantically equivalent to 'if not'. However, some authors have claimed that 'unless' is more closely related to 'only if' than to 'if not'. We report two experiments that compare conditional inferences from 'unless' to 'if-not'' and 'only if' factual and counterfactual conditionals. The first experiment compared 'not-A unless B' and 'if not-B then not-A' and showed a difference between affirmative (i.e. B therefore A, A therefore B) and negative (i.e. not-B therefore not-A, not-A therefore not-B) inferences only for factual 'if not'. The second experiment compared 'not-A unless B' and 'A only if B' and showed no difference between affirmative and negative inferences for factual 'unless' and 'only if', whereas the affirmative inferences were higher for counterfactual 'unless' and 'only if'. In both experiments latency results confirmed that inferences from 'B to A' were faster than from 'A to B' for 'unless' and 'only if'. The implications of the results for the mental representations and processing of counterfactual 'unless', 'if not' and 'only if' are discussed in the context of mental model theory.

The psychology of conditional reasoning has mainly centered on the study of indicative conditionals of the form 'if A then B' (Evans, Newstead, & Byrne, 1993). Our aim in this paper is not to study standard indicative 'if

[*] Acknowledgements: We wish to thank Ruth Byrne and Suzanne Egan for their ideas and comments. This research was supported by the following grants: Spanish Ministry of Education and Science, Grant PB98-C020-001; and Portuguese Foundation for Science and Technology (FCT), POCTI/PSI/42215/2001. Some of the results were presented at the first international conference on cognitive processes and context in counterfactual thinking, Lisbon, 2002; and at the second international conference on reasoning and decision-making, Padua, 2003. *Address for correspondence:* Prof. Juan A. García-Madruga, Department of Developmental and Educational Psychology, UNED – 28023, Madrid, Spain**.** Tel: 913986259; fax: 913987951**.** Email: jmadruga@psi.uned.es

then' conditionals, but rather two particular kinds of conditionals: subjunctive counterfactuals and 'unless' formulations. 'Unless' is a negative conditional connective that, according to philosophers, is semantically equivalent to 'if not' (e.g., Quine, 1972; Reichenbach, 1947). An 'unless' expression such as 'Virginia will not pass the exam unless she works harder' is therefore supposedly equivalent to 'If Virginia does not work harder she will not pass the exam'. However, linguists and psycholinguists have often disagreed with this opinion. According to Fillenbaum (1986), an important part of the 'unless' expression, the stress on the need to 'work harder', is lost in the 'if not' expression. Fillenbaum (1986) has claimed that 'unless' is more closely related to 'only if' than to 'if not'. The illocutionary force of the 'unless' expression is sustained better by using an 'only if' formulation, such as 'Virginia will pass the exam only if she works harder'.

People often generate counterfactuals to express what might have been different in the past (Kahneman & Tversky, 1982). Counterfactual thoughts may play an important role in cognition, emotion, and social judgements (see Byrne, 2005). For example, suppose a friend just took an important exam. After learning the result of the exam, you might make a conjecture such as 'If Virginia had not worked hard she would not have passed the exam'. Counterfactual thoughts can be expressed in other ways, too. For example, you might have thought instead, 'Virginia would not have passed the exam unless she had worked hard', or even, 'Virginia would have passed the exam only if she had worked hard'. The 'if', 'unless' and 'only if' counterfactuals seem to emphasize different aspects of the causal sequence of events. However, unless counterfactuals have not been studied yet. Our aim is to compare counterfactual conditionals based on 'unless' to ones based on 'if not' and 'only if'. We wish to examine whether people think differently about counterfactuals based on these three conditional expressions, and whether they make different inferences from them.

First, we consider the view that people rely initially only on one possibility to understand 'if' and 'if not', and we review evidence that they rely on two possibilities to understand counterfactual 'if' and 'if not'. Second, we consider the view that people rely initially on two possibilities to understand 'only if' and 'unless', in both factual and counterfactual expressions. We report two experiments that compare factual and counterfactual conditionals of these three different sorts: 'unless', 'if not' and 'only if'. This article is the first to have studied 'unless' counterfactuals. Our aim is to contribute to the literature on reasoning with counterfactuals, and especially to increase our knowledge regarding human reasoning with 'unless' conditionals.

Factual and Counterfactual 'If' and 'If not'

An 'if A then B' conditional, such as 'if Cristina goes to the conference, then she takes her laptop' is logically equivalent to 'if not-B then not-A', as in 'if Cristina does not take her laptop, then she does not go to the conference'. An 'if A then B' conditional is false when the antecedent is true and the consequent false, Cristina goes to the conference and she does not take her laptop, 'A and not-B'. It is obviously true when its antecedent and consequent are true, Cristina goes to the conference and she takes her laptop, that is 'A and B'. It is also true when the antecedent and consequent are false, Cristina does not go to the conference and she does not take her laptop, 'not-A and not-B'. Finally, consider the situation in which the antecedent is false and the consequent true, Cristina does not go to the conference but she takes her laptop, 'not-A and B'. Given a 'conditional' interpretation it is true (Cristina does not go to the conference but she takes her laptop perhaps because she is on holidays), but given a 'biconditional' intepretation, it is false. People readily come to different interpretations of conditionals (Byrne, Espino, & Santamaría, 1999; Johnson-Laird & Byrne, 2002). As can be observed in Table 1, the truth conditions for 'if A then B' and 'if not-B then not-A' are the same.

Table 1: The truth conditions and four inferences for 'if Cristina goes to the conference then she takes her laptop' and 'if Cristina does not take her laptop then she does not go to the conference '.

Truth conditions		
She goes to the conference and she takes her laptop		True
She does not go to the conference and she does not take her laptop		True
She does not go to the conference and she takes her laptop		True
She goes to the conference and she does not take her laptop		False
Inferences		
'If A then B'		*'If not-B then not-A'*
MP	She goes to the conference, therefore she takes her laptop	MT
AC	She takes her laptop, therefore she goes to the conference	DA
MT	She does not take her laptop, therefore she does not go to the conference	MP
DA	She does not go to the conference, therefore she does not take her laptop	AC

There are different theoretical views on conditional reasoning. One view is that reasoners rely on abstract rules of inference that operate in virtue of their form (e.g., Braine & O'Brien, 1998; Rips, 1994), and a second view is that reasoners rely on domain-specific rules of inference (e.g., Fiddick, Cosmides & Tooby, 2000; Holyoak & Cheng, 1995). A third view is that reasoners rely on imagining possibilities or mental models (Johnson-Laird & Byrne, 2002). A fourth view claims that reasoners deal not only with possibilities as mental models, but also attach knowledge and probabilities to these models; in this way, these new kind of models are not semantic but 'epistemic' models (Evans & Over, 2004; Evans, Over & Handley, 2005). The mental model theory is the only view that has offered a corroborated account of reasoning with counterfactual conditionals, and so we will outline it further here (e.g., Byrne, 2005; Byrne & Tasso, 1999; Thompson & Byrne, 2002; Santamaría, Espino & Byrne, 2005).

According to the mental model theory, people understand and reason about 'if' by relying on particular possibilities. A main assumption is that people keep in mind few possibilities at the outset because of constraints their limited working memories place on them (Johnson-Laird & Byrne, 1991; Johnson-Laird, Byrne, & Schaeken, 1992). In other words, for 'if A then B' they keep in mind just the possibility explicitly mentioned in the conditional, 'A and B'.

In this way, the information needed to make the affirmative inferences (MP: A therefore B, and AC: B therefore A) corresponds to reasoners' initial understanding, i.e., to the affirmative possibility, 'A and B'. Hence affirmative inferences can be made readily. The information necessary for the negative inferences (MT: not-B therefore not-A, and DA: not-A therefore not-B), on the other hand, does not correspond to the initial true possibility. Instead, reasoners must 'flesh out' their understanding of the conditional in order to think about other true possibilities, such as 'not-A and not-B'. Once they have thought about the negative possibility they are able to make the negative inferences.

Although originally the theory of mental models proposed that 'if not' negated conditionals were likely to elicit both a negated representation and a corresponding non-negated one (see Johnson-Laird and Byrne, 1991), more recently Evans and collaborators have refuted this idea (see Evans, 1993; Evans, Clibbens and Rood, 1996; Evans and Handley, 1999; Evans, Legrenzi and Girotto, 1998). Accordingly, the representation of negated and non-negated conditionals is based on the same common principle of truth: Mental models represent true assertions whether negative or affirmative. Thus for 'if not-B then not-A', reasoners keep in mind a

single possibility (not-B and not-A), just as they do for 'if A then B' (A and B) (see Johnson-Laird and Byrne, 1992, 2002).

Table 2: The initial possibilities reasoners keep in mind when they interpret factual and counterfactual conditionals for 'unless', 'if not' and 'only if'.

	Factual	Counterfactual	
Not-A unless B	not-B and not-A B and A	Fact Counterfactual:	B and A not-B and not-A
If not-B then not-A	not-B and not-A	Fact Counterfactual:	B and A not-B and not-A
A only if B	B and A not-B and not-A	Fact Counterfactual:	not-B and not-A B and A

Note: The full set of true possibilities for each of the three linguistic forms is the same: 'A and B', 'not-A and not-B', and 'not-A and B'.

The mental model account also extends to counterfactual conditionals, which have been relatively neglected in psychology until recently (e.g., Byrne & Tasso, 1999; Thompson & Byrne, 2002; Revlin, Cate & Rouss, 2001; Evans & Over, 2004). The idea is that an 'if not-B then not-A' counterfactual implies something more than just that stated in an explicit factual. Consider the sentence 'if Cristina had not taken her laptop, then she would not have gone to the conference'. This sentence implicitly seems to imply that Cristina did take her laptop and she did go to the conference. Attempts to develop a clear account of the meaning of counterfactuals led to the development of possible world semantics in philosophy (e.g., Stalnaker, 1968; Lewis, 1973; Pollack, 1986; Jackson, 1987). The subjunctive mood the counterfactual, e.g., 'if B had not been the case then A would not have been the case' helps convey the presupposition that in fact, 'B and A' is the case (Fillenbaum, 1974).

According to the mental model theory, people may understand a counterfactual by keeping in mind the explicit conjecture, 'not-B and not-A', and also the presupposed facts, 'B and A', and they may keep track of

the epistemic status of these possibilities as factual or counterfactual (Johnson-Laird & Byrne, 1991), as Table 2 shows. As a result of keeping these two possibilities in mind, reasoners make different inferences from counterfactuals and factuals. They make more affirmative inferences (A therefore B, and B therefore A) from the counterfactual because they also have access to the affirmative possibility counterfactual implies (Byrne & Tasso, 1999). They make the same frequency of negative inferences (not-B therefore not-A, and not-A therefore not-B) because they have access to the negative possibility for both the counterfactual and the factual.

On the other hand, more recently Evans & Over (2004) have claimed that the comprehension of counterfactual conditionals leads to the construction of a single mental model or possibility. The representation that people build when they understand a factual indicative such as 'if Cristina takes her laptop then she goes to the conference' and a counterfactual conditional, as 'if Cristina had not taken her laptop, then she would not have gone to the conference', is the same. This single model includes 'semantic' information -and also 'pragmatic' knowledge and implicatures. The increase of affirmative inferences (A therefore B, and B therefore A) in counterfactuals is the result of the pragmatic knowledge attached to the epistemic model for counterfactuals; in other words, these inferences would be pragmatic, not logical.

Factual and Counterfactual 'Unless' and 'Only if'.

Not-A unless B' is equivalent to 'if not-B then not-A' (e.g., Quine, 1972; Reichenbach, 1947). The equivalence of, for example, 'Cristina does not go to the conference unless she takes her laptop' to 'if Cristina does not take her laptop then she does not go to the conference ' can be easily appreciated, the same situation renders both sentences false: Cristina goes to the conference and she does not take her laptop. 'Not-A unless B' is false in the situation 'A and not-B' and true otherwise, that is, in the situations 'A and B', 'not-A and not-B', and 'not-A and B' (see Table 3, column 1).

Psychologically, however, 'Not-A unless B' and 'if not-B then not-A' may be interpreted quite differently. In fact 'Not-A unless B' may also be closely related to 'Not-A except if B' (Geis, 1973; von Fintel, 1991) and to 'A only if B' (Fillenbaum, 1986; see also Wright and Hull, 1986, 1988). It has been assumed by linguists that counterfactual forms clash with 'unless' and are unacceptable (e.g. Geis, 1973; Dancygier, 1998). However, more recently it has been demonstrated that 'unless' counterfactual sentences, although probably rare, are used in natural discourse and are not in any way

unacceptable (see Declerck and Reed, 2000; Dancygier, 2002). There have been few psychological studies of factual 'unless' (Fillenbaum, 1986; Wright and Hull, 1986; 1988) and there is not yet in our knowledge any psychological study on counterfactual 'unless'. Studies on factual 'unless' have established that reasoners make fewer inferences from 'not-A unless B' compared to 'A only if B', and to 'if not-B then not-A' (Carriedo, García-Madruga, Moreno & Gutiérrez, 1999; García-Madruga, Gutiérrez, Carriedo, Moreno, & Johnson-Laird, 2002; see also García-Madruga, Carriedo, Moreno, Gutiérrez & Schaeken, 2008; Schaeken, García-Madruga & D'Ydewalle, 1997).

Table 3: The truth conditions for the three linguistic forms and the two kind of inferences from them.

	Not-A unless B	A only if B	If not-B not-A
Truth conditions			
A and B	True	True	True
not-A and not-B	True	True	True
not-A and B	True	True	True
A and not-B	False	False	False
Affirmative Inferences			
	A.˙.B (MP)	A.˙.B (MP)	A.˙.B (MT)
	B.˙.A (AC)	B.˙.A (AC)	B.˙.A (DA)
Negative Inferences			
	not-B.˙.not-A (MT)	not-B.˙.not-A (MT)	not-B.˙.not-A (MP)
	not-A.˙.not-B (DA)	not-A.˙.not-B (DA)	not-A.˙.not-B (AC)

The special difficulty of 'unless' conditionals comes probably from its typical 'asymmetric' formulation: 'Not-A unless B': the antecedent is negated but the consequent is not. What, if anything, follows from 'Not-A unless B' and 'Not-A'? Recall that the major premise is true in the situations 'A and B', 'not-A and not-B', and 'not-A and B'. Hence, when you know that 'not-A' is the case, you do not know whether 'B' or 'not-B' is the case. We will refer to this inference as a DA for comparison with the other

linguistic forms[1]. Given 'not-A unless B', and 'not-A' the DA inference is to conclude 'not-B', but some reasoners tend to infer 'B' instead. In fact, they tend to make such asymmetric responses for each of the inferences from 'unless'. The tendency may reflect a superficial strategy based on matching the minor premise (not-A) and conclusion (B) with the elements in the major premise (not-A unless B) (García-Madruga et al, 2002; Schaeken, et al, 1997).

Following Geis's (1973) ideas, psychological studies have shown that 'Not-A unless B' may be closely related to 'A only if B' (Fillenbaum, 1986; see also García Madruga et al, 2002; Wright and Hull, 1986, 1988). 'Cristina goes to the conference only if she takes her laptop' is logically equivalent to 'Cristina does not go to the conference unless she takes her laptop', that is the same situation renders both sentences false: Cristina goes to the conference and she does not take her laptop. 'A only if B' is false in the situation 'A and not-B', and true in the situations 'A and B', 'not-A and not-B', and 'not-A and B' (see Table 3, column 2).

There is a considerable debate about the representation and inferences from 'only if' conditionals. Three main hypotheses have been proposed about the mental representation of 'only if': First, reasoners keep two possibilities in mind for 'A only if B': 'A and B' and 'not-A and not-B' (Johnson-Laird & Byrne, 1989); as a result, they can readily make both MP (A therefore B) and MT (not-B therefore not-A). This hypothesis accounts for the main result for 'only if' conditionals: the classical difference between MP and MT inferences for 'if A then B' tend to disappear. Reasoners can make both inferences readily (e.g., Evans & Beck, 1981; Roberge, 1978). However, according to this account, they should make more DA (not-A therefore not-B) from 'only if', but the same frequency of AC (B therefore A) from 'only if' and 'if', and they do not.

The second hypothesis claims that, as in 'if A then B', reasoners keep a single possibility in mind for 'A only if B', but in the direction 'B and A'. Therefore, reasoners have a processing preference for making inferences from B to A (Espino & Hernández, 2009; Evans, 1993; Santamaría &

[1] Naming the four inferences for 'unless' is difficult and we have been guided by the form of the minor premise and conclusion, as well as by the possibilities that are true and false. As Table 3 shows, for 'Not-A unless B', the three true possibilities are 'A and B', 'not-A and not-B' and 'not-A and B'. The two logically valid inferences are 'A therefore B', and 'not-B therefore not-A'. The two fallacies are 'not-A therefore not-B', and 'B therefore A'. We call these inferences MP, MT, DA and AC respectively, although we accept that at first sight it may seem odd to use 'MP' to refer to the inference 'Not-A unless B, A therefore B'. Other naming schemes run into difficulties however and so we opt for this one.

Espino, 2002; Grosset & Barrouillet, 2003; see however Oberauer, Hornig, Weidenfeld & Wilhelm, 2005; Oberauer & Wilhelm, 2000). As a result, the backward inferences AC (B therefore A) and MT (not-B therefore not-A) are made more often from 'A only if B' (e.g., Evans, 1993; Evans, Clibbens & Rood, 1995; 1996). However, on this account reasoners should make fewer MT (not-B therefore not-A) than AC (B therefore A) from 'only if', and they do not.

A third and more recent hypothesis about the mental representation of 'only if' combines aspects of the two previous hypotheses: Reasoners understand 'A only if B' by thinking about two possibilities, but in a directional manner: 'B and A' and 'not-B and not-A' (Carriedo et al, 1999; Egan, García-Madruga & Byrne, 2008; García-Madruga et al, 2002; see Table 2). This hypothesis may account for main empirical results on 'only if' conditionals, namely the lack of differences between MP (A therefore B) and MT (not-B therefore not-A) and the processing preference for making inferences from B to A.

The first study on counterfactual 'only if' expressions by Egan and collaborators (2008) have tested this third hypothesis about the mental representation and processing of 'only if'. These authors compared factual and counterfactual 'only if', e.g., 'Cristina would have gone to the conference only if she had taken her laptop'. On their account factual and counterfactual 'only if' are both understood by keeping in mind two possibilities. For factual 'only if' the two possibilities are true possibilities, whereas for counterfactual 'only if', one of these possibilities is understood as a fact ('not-B and not-A'), and the other is understood as a counterfactual conjecture ('B and A', see Table 2). In this paper we extend the account of counterfactual 'only if' to counterfactual 'unless', e.g. 'Cristina would not have gone to the conference unless she had taken her laptop'. According to our view, factual and counterfactual 'unless' are understood by keeping also in mind two possibilities. The two possibilities are true possibilities for factual 'unless', whereas for counterfactual 'unless', one possibility is factual ('B and A') and the other counterfactual ('not-B and not-A', see Table 2).

In this paper we will test our account about representation and processing of 'unless' factual and counterfactuals by comparing them to 'if-not'' and 'only if' conditionals. In experiment 1 we will compare factual and counterfactual 'not-A unless B' and 'if not-B then not-A'. According to our view reasoners keep in mind only one possibility for factual 'if not' but two possibilities for counterfactual 'if not' and factual and counterfactual 'unless'. In experiment 2 we will compare factual and counterfactual 'not-A

unless B' and 'A only if B'. According to our view reasoners keep in mind two possibilities for factual and counterfactual 'unless' and 'only if'.

EXPERIMENT 1

'Not-A unless B' and 'If not-B then not-A' Counterfactuals

The aim of our first experiment was to compare factual and counterfactual 'if not ' to factual and counterfactual 'unless'. We measured the frequency of the affirmative and negative inferences. The inferences for 'not-A unless B' and 'if not-B then not-A' are shown in Table 3. We propose that reasoners keep in mind different possibilities to understand 'Not A unless B' (not-B and not-A, B and A) compared to 'if not-B then not-B' (not-B and not-A). We also propose that reasoners keep in mind different possibilities for factual or counterfactual conditionals (a single possibility compared to two possibilities). As a result, we expect reasoners given 'if not-B then not-A' to make more affirmative inferences (A therefore B, and B therefore A) from the counterfactual than from the factual conditional (see table 2). In contrast, we expect to find that reasoners make the same frequency of affirmative and negative inferences from counterfactual and factual 'unless'.

Time measures show a special sensitivity to directional effects in the representation. Grosset & Barrouillet (2003; see also Espino & Santamaría, 2002) showed that reasoners were faster drawing backward 'B to A' for 'A only if B' in comparison with faster forward 'A to B' inferences for 'if A then B'. This backward directional effect for 'only if' has been also extended to 'unless' (García-Madruga et al., 2008). The predicted directional effect for 'if not-B then not-A' is from 'B to A', although in this case it is a forward effect. The directional effect is included in our proposed representation (see Table 2) and thus we predict that latencies from 'B to A' will be faster for 'if not-B then not-A' and 'Not-A unless B'. We measured latencies from 'A to B' inferences and from 'B to A' inferences.

Summarizing, our predictions are:

1. A higher percentage of endorsements for 'if-not' than for 'unless'. Likewise, we expect a higher percentage of 'asymmetric' responses for 'unless' than 'if not'.

2. A higher percentage of negative inferences than affirmative inferences for factual 'if-not'. No differences between affirmative and negative inferences for factual and counterfactual 'unless', and for counterfactual 'if not'.

3. Lower latencies for inferences in the 'B to A' direction than inferences in the 'A to B' direction, for both 'unless' and 'if not' factual and counterfactual conditionals.

METHOD

Participants. The participants were 90 undergraduate students registered for a children's teacher degree course in Granada University, who were taking part in a motor development class. There were 35 women and 55 men and their average age was 20 years, ranging from 19 to 24. They participated voluntarily in return for credits, they had not been trained in formal logic, nor had they participated previously in any reasoning study. They were assigned at random to the 'unless' condition or the 'if-not' condition (n = 45 in each group).

Materials and Design. We constructed two sets of problems, an 'unless' set and an 'if not' set. Each set contained 32 problems, 16 based on factual conditionals and 16 based on counterfactual conditionals. The problems were based on a locations neutral content (e.g., if Victor is in Madrid then Andrés is in Sevilla). They were presented in the participants native Spanish and referred to common Spanish proper names and well-known Spanish cities. The factual conditionals were in the present tense and the counterfactuals in the past tense, and the indicative and subjunctive mood was used to convey factuality and counterfactuality as appropriate in Spanish, as illustrated in Appendix A. Each problem consisted of a conditional premise and a categorical premise corresponding to A therefore B, B therefore A, not-A therefore not-B, and not-B therefore not-A (see Table 3). Participants were given a selection of three possible conclusions (in the present tense for factuals and the past tense for counterfactuals). An example of a counterfactual problem is as follows: 1) Victor was in Madrid, 2) Victor was not in Madrid, 3) There is no valid conclusion. For half of the problems the affirmed component (Victor was in Madrid) appeared first associated with the '1' response key and for the other half it appeared second associated with the '2' key. Participants carried out 4 instances of each type of inference for each type of conditional, factual and counterfactual (i.e., 4 inferences x 2 conditional types x 4 instances = 32 problems). The content for the 32 problems for each participant was drawn at random from a pool of 64 conditionals. Each participant was given the 32 problems in a different random order.

Procedure. The participants were tested in several large groups, and the problems were presented on desktop PCs using E-Prime. The participants were given some instructions illustrated with an example based on a conjunction to familiarize themselves with the task presentation and keyboard response options. Participants were advised that they could take as long as they needed to complete the task and that they were being timed. They were told that the problems concerned drivers for a transport company, and the cities where each of the drivers may or may not be, and that during waiting times, the drivers used this type of problem as a kind of game, to allow them to know where their colleagues were. Participants were requested to consider that the premises were true, and the conclusion had to be the only possible conclusion; otherwise they should select the 'there is no valid conclusion' option. Participants pressed the space bar to view each new piece of information (the conditional, the minor premise, the conclusion set), and each remained on screen to be joined by the subsequent information. The participants pressed one of the keys labelled '1' '2' or '3' to select a conclusion, these keys were in the center of the keyboard and corresponded to the T, G, & B keys.

RESULTS AND DISCUSSION

Endorsements

We analysed the results in an ANOVA carried out on the endorsements of inferences with the factors form (if-not, unless), mood (factual, counterfactual), and inference (affirmative, negative), with repeated measures on the last two factors. It showed a main effect of form ($F(1,88)=3.9$; $Mse=10.5$; $p = .05$), and inference ($F(1,88) = 8.4$; $Mse= 2.9$; $p < .01$), but not of mood ($F(1,88)=2.3$; $Mse=1.7$; $p >.1$). Inference interacted with form and mood ($F(1,88) = 9.5$; $Mse= 2.9$; $p < .01$; $F(1,88)=1.8$; $Mse=12.0$; $p < .001$; respectively). Form and mood did not interact ($F(1,88)=0.1$; $Mse=1.7$; $p = .75$). The three-way interaction was not reliable ($F(1,88) =2.2$; $Mse= 1.8$; $p = >.05$). The percentages of endorsements can be observed in table 4.

As the significant effect of form shows, our first prediction concerning the comparison of 'if not-B then not-A' and 'Not-A unless B' conditionals was confirmed. Reasoners made fewer inferences from 'unless' (68%) than from 'if not' (76%). These results are consistent with previous experiments that shows the special difficulty of 'unless' conditionals (see García-Madruga et al, 2002; Schaeken et al, 1997). As found in previous studies, participants tended to make asymmetric conclusions, e.g., from 'not-A', to infer 'therefore B'. Such conclusions for the two kind of AC

and DA inferences were made more often for 'unless' than for 'if not' (27% and 14%; F(1,88)=12.4; MSe=4.3; p<.001; see also García-Madruga et al, 2002; Schaeken et al, 1997). Likewise, they were more frequent for factual than for counterfactual 'unless' (30% and 25%; F(1,44)=6.4; MSe=.7; p<.05). The results lend support to the idea that there may be a superficial matching strategy in operation for 'unless' (Schaeken et al, 1997).

Table 4: Percentages of conclusions endorsed by participants in Experiments 1 and 2.

	Affirmative Model			Negative Model		
EXP. 1		(A.˙.B;	B.˙.A)		(not-A.˙.not-B;	not-B.˙.not-A)
Not-A unless B						
Factual	65	(63	66)	68	(64	72)
Counterfactual	71	(71	72)	68	(64	71)
If not-B not-A						
Factual	64	(64	64)	86	(82	91)
Counterfactual	75	(77	73)	80	(75	84)
EXP. 2						
Not-A unless B						
Factual	68	(63	73)	65	(59	70)
Counterfactual	75	(74	76)	68	(66	70)
A only if B						
Factual	88	(85	90)	83	(73	92)
Counterfactual	90	(90	89)	84	(79	89)

Our second prediction for 'if not' and for 'unless' was also confirmed. There were differences in the frequency of negative and affirmative inferences for 'if not' (F(1,44) = 12.8; MSe=4.1, p < .001) and as expected this was confirmed for factual (86%vs 64%; F(1,44) = 21.7; MSe=143.9, p < .001), but not for counterfactual 'if not' (80% vs 75%; F(1,44) = 1.1; MSe=126.3, p >.2). Endorsement differences in the frequency of affirmative and negative inferences for 'unless' were not significant (F(1,44) = 1,5; MSe=2, p > .05). This result was confirmed for counterfactual and for factual 'unless' (F(1,44) = 1.3; MSe=71.8, p > .2; F(1,44) = 0.8; MSe=73.2, p > .3, respectively).

The results for 'if not' corroborate our suggestion that a single possibility is available for factual 'if not' and two possibilities for counterfactual 'if not'. People endorsed more negative than affirmative inferences from factual 'if not-B then not-A', whereas for counterfactual 'if not-B then not-A' there were no differences between affirmative and negative inferences (see also Thompson & Byrne, 2002).

The results for 'unless' also confirm our predictions. They support the idea that reasoners keep in mind two possibilities to understand factual and counterfactual 'unless'. The experiment also replicates the finding that 'unless' is a difficult conditional and that a common error for 'unless' is committing an asymmetric conclusion.

In order to check our hypothesis that people keep in mind two possibilities to understand factual and counterfactual 'unless' we have compared them with factual and counterfactual 'if not'. Although 'unless' and 'if not' factual conditionals are logically equivalent they are psychologically diverse: Reasoners keep in mind a single possibility to understand factual 'if not' and two possibilities to understand factual 'unless'. On the contrary, both type of conditionals seem to demand the construction of two possibilities in counterfactual formulation.

Latencies

The latencies were log-transformed to the base e (we removed outliers greater or less than three standard deviations of the reading times for the conditional sentences and premise-plus-conclusion response time: the overall percentage of removed scores was of 3.4%. A 2 (if not, unless) by 2 (factual, counterfactual) by 2 (MP & DA, AC & MT) ANOVA, with repeated measures on the second two factors, was carried out to check our hypothesis regarding the faster backward latencies (AC and MT) in comparison with forward latencies (MP and DA) for 'if not' and 'unless'. The ANOVA showed a main effect of directionality, participants take less time to make inferences from 'B to A' than from 'A to B' ($F(1,88)=15.7$; $Mse=1.9$; $p<.001$). There was also a main effect of mood ($F(1,88)=13.4$; $Mse=1.2$; $p<.001$); participants take less time for factual than for counterfactual inferences. The main effect of form was not reliable ($F(1,88)=8.4$; $Mse=.7$; $p<.39$). Mood did not interact with form ($F(1,88)=1.5$; $Mse=1.2$; $p<.24$), nor with directionality ($F(1,88)=1.1$; $Mse=1.9$; $p<.31$), and form and directionality did not interact ($F(1,59)=1.3$; $Mse=1.8$; $p<.27$). The three way interaction was not reliable ($F(1,88)=.13$; $Mse=1.8.$; $p<.72$). The directional effect was reliable for factual and counterfactual 'if not' ($F(1,44) = 7.3$; $MSe=2.7$, $p < .01$; $F(1,44) = 5.2$;

MSe=1.2, $p < .05$; respectively). The directional effect was marginally significant for factual 'unless' ($F(1,44) = 3.6$; MSe=1.8, $p = .06$), but for counterfactual 'unless' it was not reliable ($F(1,44) = 1.4$; MSe=1.7, $p > .05$).

The results show that people take more time to endorse inferences from counterfactual than from factual conditionals. Likewise, people take less time to endorse inferences from 'B to A' than from 'A to B' inferences. This directional effect is robust for factual 'if not', less robust for counterfactual 'if not' and factual 'unless', and not reliable for counterfactual 'unless'. This pattern suggest a possible explanation, that is, keeping in mind two possibilities may decrease the directional effect.

Table 5: Mean of the log-transformed latencies (in seconds) for the conclusions endorsed in Exp. 1 and 2.

	'B to A' inferences (B∴A, not-B∴not-A)	'A to B' inferences (A∴B, not-A∴not-B)
EXP. 1		
Not-A unless B		
Factual	5.03	5.57
Counterfactual	5.70	6.03
If not-B not-A		
Factual	4.70	5.65
Counterfactual	5.20	5.72
EXP. 2		
Not-A unless B		
Factual	8.82	8.93
Counterfactual	8.93	8.96
A only if B		
Factual	8.89	8.98
Counterfactual	8.94	9.05

EXPERIMENT 2

'A only if B' and 'Not-A unless B' Counterfactuals

The objective of this experiment was to compare 'only if' and 'unless' factual and counterfactual conditionals, in order to test the close similarity between these expressions. We measured the frequency of the affirmative and negative inferences and the latencies of inferences from 'B to A' and from 'A to B'. Granted the special difficulty of 'unless' expressions, we expect a lower percentage of inferences and a higher percentage of 'asymmetric' responses for 'unless' than for 'only if'. Since two possibilities are kept in mind for both factual and counterfactual 'only if' and 'unless', we expect the same frequency of affirmative and negative inferences from factual and counterfactual 'only if', and for factual and counterfactual 'unless' (see Egan et al, 2008) Granted the backward directional representation of 'only if' and 'unless', we also expect lower latencies for inferences from 'B to A' than from 'A to B', for both 'unless' and 'only if' factual and counterfactual conditionals.

Summarizing, our predictions are:

1. A higher percentage of endorsements for 'only if' than for 'unless' and a higher percentage of 'asymmetric' responses for 'unless' than for 'if not'.

2. No differences between affirmative and negative inferences for factual and counterfactual 'unless' and 'only if'.

3. Lower latencies for inferences from 'B to A' than from 'A to B' for both 'unless' and 'only if' factual and counterfactual conditionals.

METHOD

Participants. The participants were 114 undergraduate psychology students (81 women and 33 men, with an average age of 21 years) from Lisbon's Instituto Superior de Psicologia Aplicada. They participated voluntarily in the experiment, were not trained in formal logic, and had not previously participated in any experiment on reasoning. They were assigned at random to the 'only if' condition (n=55) or the 'unless' condition (n=59).

Materials, Design and Procedure. The materials were the same as the experiment 1 except that now we compare 'unless' with 'only if'. We constructed two sets of problems, an 'unless' set and an 'only if' set. Each set contained 32 problems, 16 based on factual conditionals and 16 based on

counterfactual conditionals. The problems were given to the participants in their native Portuguese and we assigned at random 32 different contents to the problems based on common Portuguese proper names and well known Portuguese cities. The design and procedure was the same as the previous experiment.

RESULTS AND DISCUSSION

Endorsements

We analysed the results in an ANOVA carried out on the endorsements of conclusions with the factors form (only if, unless), mood (factual, counterfactual), and inference (affirmative, negative) with repeated measures on the second two factors. As expected, the main effect of form was reliable ($F(1,112)=21.17$; $Mse=.16$; $p < .00$). The main effects of mood and inference were also reliable; participants made more inferences for counterfactual than for factual (79% versus 74%; $F(1,112)=5.51$; $Mse=0.02$; $p < .02$) and they made more affirmative than negative inferences (80% versus 75%; $F(1,112)=11.47$; $Mse=0.02$; $p < .00$). Mood did not interact with inference ($F(1,112)=0.93$; $Mse=0.02$; $p < .34$), or with form ($F(1,112)=1.55$; $Mse=0.02$; $p<.22$), and form and inference did not interact ($F(1,112)=0.00$; $Mse=0.02$; $p< 0.99$). The three-way interaction was not reliable ($F (1,112) = 0.33$; $Mse=0.02$; $p< 0.57$). The percentages of endorsements can be observed in table 4.

Our first prediction concerning the comparison of 'A only if B' and 'Not-A unless B' conditionals was confirmed; participants made more inferences for 'only if' than for 'unless' (87% versus 69%). These results are consistent with previous experiments that shows the special difficulty of 'unless' conditionals (see, García-Madruga et al, 2002; Schaeken et al, 1997). As found in previous studies, participants tended to make asymmetric conclusions, e.g., from 'not-A', to infer 'therefore B'. Such conclusions for the two kind of inferences were made more often for 'unless' than for 'only if' (22 % vs 3% $F (1,112)=35.43$; $Mse=0.96$; $p<0.00$, see also García-Madruga et al, 2002; Schaeken et al, 1997). Participants also made these errors more often for factual than counterfactual 'unless' (26% vs 19% , $F (1,112)=5.08$; $Mse=0.43$; $p< .03$). These results confirm those found in Experiment 1.

As for our second prediction, there was a difference in the frequency of affirmative and negative inferences for 'only if' ($F (1,54)=5,57$; $Mse=0.02$; $p< .02$) and this result was confirmed for counterfactual only if, for which people endorsed more affirmative than negative inferences, but not for factual 'only if' conditionals ($F (1,54)=4.41$; $Mse=.02$; $p<.04$; and

F(1,54)=2.61; Mse=0.02; p<.11, respectively. Differences in the frequency of endorsements of affirmative and negative inferences for 'unless' were also significant (F(1,58)=5.92; Mse=.02; p<.02). This result was confirmed for counterfactual conditionals (F (1,58)=7.30; Mse=.02; p<.01) for which people endorsed more affirmative than negative inferences, but not for factual 'unless' (F (1,58)=1.37; Mse=.02; p<.24).

The results for 'only if' do not corroborate our suggestion that the same two possibilities are available for factual and counterfactual 'only if'. Reasoners endorsed the same amount of negative and affirmative inferences for factual 'only if' suggesting that they keep in mind two possibilities. But for counterfactual 'only if', people endorsed more affirmative inferences than affirmative inferences.

The results for 'unless' replicate only partially those of experiment 1. People endorsed the same amount of negative and affirmative inferences for factual 'unless', suggesting that they keep in mind two possibilities. However, for counterfactual 'unless', as for counterfactual 'only if', people endorsed more affirmative inferences than negative inferences. The experiment again confirms the finding that 'unless' is a difficult conditional and that a common error for 'unless' is committing an asymmetric conclusion.

The results of experiment show an unpredicted difference between affirmative and negative inferences for both sorts of counterfactual conditionals, 'only if' and 'unless'. A similar pattern of differences between affirmative and negative inferences was also found by Egan et al (2008) for counterfactual 'only if'. We will analyse it in a deeper way in the General Discussion.

Latencies

The latencies were log-transformed to the base e (we removed outliers in the same way as in exp. 1: the overall percentage of removed scores was of 10.7%). A 2 (only if, unless) by 2 (factual, counterfactual) by 2 (MP & DA, AC & MT) ANOVA, with repeated measures on the second two factors, was carried out to check our hypothesis about the faster backward latencies (AC and DA) in comparison with forward latencies (MP and DA) for 'only if' and 'unless'. The ANOVA showed a main effect of directionality; participants take less time to make backward from 'B to A' inferences than forward from 'A to B' inferences (F(1,112)=22.20 ; Mse=.06 ; p<.00). There was also a marginal main effect of mood (F(1, 112)=3.50; Mse=.04; p<.07); participants take less time for factual than for

counterfactual inferences. The main effect of form was not reliable ($F(1, 112)=0.50$; $Mse=.25$; $p<.48$). Mood did not interact with form ($F(1,88)=1.20$; $Mse=.04$; $p<.27$), nor with directionality ($F(1, 112)=.1.60$; $Mse=.05$; $p<.21$) and form and directionality did not interact ($F(1, 112)=1.10$; $Mse=.06$; $p<.29$). The three way interaction was reliable ($F(1, 112)=5.90$; $Mse=.05.$; $p<.02$). This directional effect was not reliable for factual but it was for counterfactual 'only if' ($F(1,54)=1.89$; $Mse=.06$; $p<.17$; and $F(1, 54)=6.41$; $Mse=.06$; $p<.01$; respectively). However, it was reliable for factual but not for counterfactual 'unless' ($F(1, 58)=21.51$; $Mse=.06$; $p<.00$; and $F(1, 58)=1.49$ $Mse=.07$; $p<.22$; respectively).

The results replicate the main findings of experiment 1.: People take more time to endorse inferences from counterfactual than from factual conditionals. Likewise, people tended to take less time to endorse inferences from 'B to A' than from 'A to B' inferences. However, this directional effect is not robust neither for 'unless' nor for 'only if': the difference was not reliable for factual 'only if' and for counterfactual 'unless' conditionals. Latency results of this experiment are higher than those of experiment 1. These differences may have diverse origins although they are probably due, at least partially, to the linguistic differences between Portuguese and Spanish; for instance, a counterfactual 'unless' expression in Spanish contains 14 words whereas in Portuguese has 17 words.

GENERAL DISCUSSION

The main result of these experiments is that reasoners are able to understand and reason from counterfactual 'unless' assertions. This predicted result rebut traditional linguists ideas on the unacceptability of 'unless' counterfactuals (e.g. Geis, 1973; Dancygier, 1998) and confirms more recent linguistic ideas (see Declerck and Reed, 2000; Dancygier, 2002). Likewise, our experiments confirm that reasoners trait counterfactual 'unless' expressions in similar way to factual 'unless', but not identically: We have found a trend to increase the difference between affirmative and negative inferences in counterfactual 'unless' and 'only if'. This unpredicted result deserves a deeper analysis, but there are other interesting results that have to be analysed before.

The results of experiment 1 confirm the mental model theory of factual and counterfactual conditionals. For factual 'if not B then not-A' conditionals, people have in mind from the outset only a single possibility corresponding to not-B and not-A (Johnson-Laird & Byrne, 2002). However, as Fillenbaum (1974) proposed, counterfactual formulations

introduce a change in conditionals. The subjunctive mood helps to convey a factual presupposition that negates the linguistically expressed possibility. The results of exp.1 for 'if not B then not A' counterfactuals suggest that people keep in mind from the outset not only the counterfactual possibility (not-B and not-A) but also the presupposed possibility (B and A). The result for 'if not-B then not-A' counterfactuals replicates and extends earlier studies (see, Byrne and Tasso, 1999; Johnson-Laird and Byrne, 1991, 2002; Thompson and Byrne, 2002).

The results of exp. 1 and 2 for 'unless' confirm prior findings about the special difficulty of 'unless' expressions. Inferential endorsements are lower and asymmetrical responses are higher for 'unless' than both for 'if not' and 'only if' (Carriedo et al, 1999; García-Madruga et al, 2002; see also Schaeken, García-Madruga & D'Ydewalle, 1997). The asymmetrical responses imply a complete misunderstanding of the meaning of 'unless' and may likely be result of a sort of superficial bias or strategy. García Madruga and colleagues (García-Madruga, Gutiérrez, Carriedo, Luzón & Vila , 2005; García-Madruga, Gutiérrez, Carriedo, Vila & Luzón, 2007) have found that asymmetrical responses are given less by high working memory subjects than by low working memory subjects, thereby showing its superficial and heuristic nature. High working memory subjects are able to resist and inhibit these kinds of responses. The decrease in asymmetrical responses in counterfactual 'unless' shows that the counterfactual formulation facilitates the comprehension of 'unless' inferences.

We have proposed two new suggestions about the mental representations and processes underlying 'unless' and 'only if'. Our first suggestion was that factual 'unless' and 'only if' are both understood by keeping in mind two possibilities, corresponding to the true possibility whose elements are affirmed, and the true possibility whose elements are negated. Moreover, the representation observes a backwards B-A direction rather than a forwards A-B direction. 'A only if B' is understood by keeping in mind initially 'B and A' and 'not-B and not-A'. Likewise, 'Not-A unless B' is understood by keeping in mind initially 'B and A' and 'not-B and not-A'. The absence of differences between affirmative and negative inferences in both Experiments for factual 'unless' and in Experiment 2 for factual 'only if' confirm our suggestion. Likewise, latency results in both experiments confirm that the representation is from 'B to A' although this backwards directionality in the representation tends to be less clear-cut for 'unless' and 'only if' than for 'if not'. These results replicate those found in other studies on factual 'unless' and 'only if' (Egan et al, 2008; García-Madruga et al, 2008).

Our second suggestion was that counterfactual 'unless' and 'only if' are understood by keeping in mind the same two possibilities. Counterfactual 'not-A would have been, unless B had been' is understood by keeping initially in mind the same two possibilities as factual 'not-A unless B', the conjecture 'not-B and not-A', and the facts 'B and A'. Counterfactual 'A would have been only if B had been' is understood by keeping in mind initially the same two possibilities as factual 'A only if B', the conjecture, 'B and A', and the facts 'not-B and not-A'. However, the results of exp. 2 show a small but reliable difference between affirmative (A.˙.B; B.˙.A) and negative (not-A.˙.not-B; not-B.˙.not-A) inferences in counterfactuals. In fact, the same tendency to increase the difference between affirmative and negative inferences in counterfactual 'unless' was found in exp. 1, although this difference was not reliable.

In order to explain these results we should not say that counterfactual formulations increase the availability of the factual possibilities. This cannot be so, as these factual possibilities are different for 'unless' and 'only if'. For 'A wouldn't have been the case unless B had been the case' the affirmative possibility 'B and A' is the factual one, whereas for 'A would have been the case only if B had been the case', the factual possibility is the negative 'not-B and not-A'.

A possible explanation might be that the introduction of a counterfactual formulation in 'only if' and 'unless' expressions modifies the pragmatic status of the two possibilities. Thus, the counterfactual formulation may increase the pragmatic stress on the affirmative model that characterise both 'unless' and 'only if' expressions. As we analysed above, following Fillenbaum (1976), the illocutionary force of 'unless' and 'only if' is focused on the affirmative possibility.

Thus, the following two sentences:

'Virginia will not pass the exam unless she works harder'

and

'Virginia will pass the exam only if she works harder'

share the same stress on the need to work harder in order to pass the exam (the affirmative possibility) and the conjecture that she is likely not doing so (the negative possibility). What about counterfactual 'unless' and 'only if'? The sentence 'Virginia would not have passed the exam unless she had worked harder' leads reasoners to imagine the same possibilities although now the affirmative is a fact, 'Virginia works harder and passes the exam' and the negative is contrary-to-facts, 'Virginia does not work harder and does not pass the exam'. The affirmative possibility implies so a causal relationship in which 'working harder' is a necessary condition to

pass the exam. This causal relationship is stressed as crucial explanation of the results observed.

Likewise, the sentence 'Virginia would have passed the exam only if she had worked harder' leads reasoners to imagine the negative but factual possibility, 'Virginia did not work hard and did not pass the exam', and the contrary-to-facts affirmative possibility, 'Virginia worked harder and passed the exam'. Again the affirmative possibility implies a causal relationship, or 'working harder' as necessary condition for 'passing the exam', that is stressed as crucial to undo the negative result of not having passed the exam. However, the materials used in the experiments were non-causal so this proposal can be considered only as a suggestion needed of further empirical support.

The pragmatic explanation of the small increase in the difference between affirmative and negative inferences in counterfactual 'only if' and 'unless' can be accommodated within mental model theory framework. 'Pragmatic modulation' proposed by Johnson-Laird and Byrne (2002) might explain this difference between affirmative and negative model inferences. Given the difficulty always implied by the construction and simultaneous maintenance of two models in working memory, the pragmatic factors mentioned above will increase the availability of the affirmative model in counterfactual formulations.

Might a different explanation of the difference between affirmative and negative inferences in counterfactual 'unless' and 'only if' be proposed? An obvious candidate is Evans and colleagues's theory (Evans & Over, 2004; Evans, Over & Handley, 2005). This theory claims that the meaning of conditionals are represented by means of 'epistemic' mental models that include not only the states of world but also 'states of belief and knowledge' (Evans, 2006). As we have just seen, the use of counterfactual expressions for 'unless' and 'only if' modifies reasoners' beliefs and knowledge thereby increasing the causal connection between 'working harder' and 'passing the exam'. Therefore, Evans and colleagues's theory is particularly well fit to explain our finding of a tendency to increase the difference between affirmative and negative inferences in counterfactual formulations. However, following the 'singularity principle', which claims that people construct only one possibility or mental model at a time (eg. Evans & Over, 2004), this theory runs clearly into difficulties explaining some other and more robust results of our experiments, particularly two of them: In Exp. 1, the pattern of the absence of differences between affirmative and negative inferences for counterfactual 'if not' and factual

'unless'; in Exp. 2, the similar pattern of the absence of differences between affirmative and negative inferences for factual 'unless' and 'only if'.

Our experiments shed some further light on the way people think about 'unless' and 'only if'. They confirm the idea that reasoners keep two possibilities in mind for 'A only if B' but in the direction 'B and A' (Evans, 1993; Grosset & Barrouillet, 2003; Santamaría & Espino, 2002). However, 'unless' and 'only if' are not entirely similar. 'Unless' conditionals are harder than 'only if' ones and asymmetric inferences are peculiar to 'unless' inferences. Our experiments also provide the first data on inferences from counterfactual 'unless'. Counterfactual formulations increase latencies of inferences from 'unless' and 'only if', and seems also to affect to the inferences that people draw from them. The rich complexity of language allows counterfactual thoughts of varying nuances to be conveyed readily, as our exploration of 'unless' and 'only if' counterfactuals suggests.

RESUMEN

Razonamiento con condicionales contrafácticos 'A menos que'. Este artículo aborda el estudio de enunciados contrafácticos con 'a menos que', tales como 'Virginia no aprobará el examen a menos que estudie más' y 'Virginia no habría aprobado el examen a menos que hubiera estudiado más'. 'A menos que' es un condicional negativo que es semánticamente equivalente a 'si no'; sin embargo, algunos autores han sostenido que 'a menos que' está más estrechamente relacionado con 'sólo si' que con 'si no'. En este trabajo se presentan dos experimentos en los que se comparan las inferencias condicionales a partir de enunciados fácticos y contrafácticos con 'a menos que', 'si no' y 'sólo si'. En el primer experimento se comparó 'no-A a menos que B' y 'si no-B entonces no-A' y se encontraron diferencias sólo en los enunciados 'si no' fácticos entre las inferencias afirmativas (B luego A, A luego B) y negativas (no-B luego no-A, no-A luego no-B). En el segundo experimento se comparó 'no-A a menos que B' y 'A sólo si B'. No se encontraron diferencias entre las inferencias afirmativas y negativas con los enunciados 'a menos que' y 'sólo si' fácticos, mientras que hubo más inferencias afirmativas que negativas con los enunciados 'a menos que' y 'sólo si' contrafácticos. En ambos experimentos las latencias de respuesta fueron más rápidas para las inferencias de 'B a A' que para las inferencias de 'A a B'. En la discusión se analizan, en el contexto de la teoría de los modelos mentales, las implicaciones de los resultados encontrados respecto al procesamiento y representaciones mentales de los enunciados contrafácticos con 'a menos que', 'si no' y 'sólo si'.

REFERENCES

Athanasiadou, A. and Dirven, R. (Eds.). (1997). *On conditionals again*. Amsterdam: John Benjamins.

Braine, M.D.S. (1978). On the relation between the natural logic of reasoning and standard logic. *Psychological Review,* 85, 1-21.

Braine, M.D.S., & O'Brien, D. (1998). *Mental Logic.* Mahwah, NJ: Lawrence Erlbaum Associates.

Byrne, R.M.J. (2002). Mental models and counterfactual thinking. *Trends in Cognitive Sciences.* 6, 405-445

Byrne, R.M.J. (2005). *The Rational Imagination: How People Create Counterfactual Alternatives to Reality.* Cambridge, M.A.: MIT Press.

Byrne, R.M.J., Espino, O. and Santamaría, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language,* 40, 347-373.

Byrne, R. M. J. & Tasso, A. (1999). Deductive reasoning with factual, possible and counterfactual conditionals. *Memory and Cognition*, 27, 726-740.

Carriedo, N., García-Madruga, J., Moreno, S. and Gutiérrez, F. (1999). How does content affect 'unless' conditional reasoning?. *Proceedings of the European Conference on Cognitive Science* (pp. 271-277). London: Psychological Press.

Cheng, P. & Holyoak, K. (1985). Pragmatic reasoning schemas. *Cognitive psychology*, 17, 391-416.

Declerck, R. & Reed, S. (2000). The semantics and pragmatics of *unless*. *English language and linguistics, 4,* 201-241

Dancygier, B. (1998). *Conditionals and prediction. Time, knowledge and causation in conditional constructions*. Cambridge: Cambridge University Press.

Dancygier, B. (2002). Mental space embeddings, counterfactuality and the use of *unless*. *English language and linguistics, 6,* 367-37.

Egan, S. M., García-Madruga, J. A. & Byrne, R. M. J. (2008). Factual and Counterfactual 'Only if' Conditionals. Paper submitted.

Espino, O. & Hernández, E. (2009). Efectos de direccionalidad en condicionales. *Psicológica*, 41-57.

Evans, J.Bt.B.T. (1993). The mental model theory of conditional reasoning: critical appraisal and revision. *Cognition*. 48, 1-20.

Evans, J.Bt.B.T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review*. In press

Evans, J.Bt.B.T. & Beck, M.A. (1981). Directionality and temporal factors in conditional reasoning. *Current psychological research,* 1, 111-120..

Evans, J.Bt.B.T., Clibbens, J. & Rood, B. (1995). Bias in conditional inference: implications for mental models and mental logic. *Quarterly Journal of Experimental Psychology,* 48, 644-670.

Evans, J.Bt.B.T., Clibbens, J. & Rood, B. (1996). The role of explicit and implicit negation in conditional reasoning. *Journal of Memory and Language,* 35, 392-409

Evans, J.Bt.B.T., Legrenzi, P. & Girotto, V. (1999). The influence of linguistic form on reasoning: the case of matching bias. *Quarterly Journal of Experimental Psychology,* 52, 185-216

Evans, J.Bt.B.T., Newstead, S.E. & Byrne, R.M.J. (1993). *Human reasoning: the psychology of deduction.* Hove UK: Psychology Press.

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

Evans J.Bt.B.T., Over, D. E. & Handley, S. J. (2005). Supposition, extensionality and conditionals: A critique of Johnson-Laird & Byrne (2002). *Psychologycal Review*,

Fiddick, L., Cosmides, L., & Tooby, J. (2000). No interpretation without representation: The role of domain-specific representations and inferences in the Wason selection task. *Cognition, 77*, 1-79.

Fillenbaum, S. (1974). Information amplified: memory for counterfactual conditionals. *Journal of Experimental Psychology, 102*, 44-49.

Fillenbaum, S. (1986). The use of conditionals in inducements and deterrents. In E. C. Traugott, A, ter Meulen, J.S. Reilly, & C.A. Ferguson (Eds*). On conditionals* (pp 179-195). Cambridge: Cambridge University Press.

García-Madruga, J.A., Carriedo, N., Moreno, S. and Schaeken, W. (1998). *Conditional reasoning from different formulations: "if", "only if" and "unless"*. Unpublished paper.

García-Madruga, J.A., Carriedo, N., Moreno, S., Gutiérrez, F. and Schaeken, W. (2008). Unless reasoning. *Spanish Journal of Psychology, 11(2),*386-399.

García-Madruga, J. A, Gutiérrez, F, Carriedo, N., Luzón J.M. & Vila, J. O. (2005). Working memory and propositional reasoning: Searching for new working memory tests. In V. Girotto and P.N. Johnson-Laird (eds.), *The shape of reason. Essays in honour of Paolo Legrenzi* (pp. 69-89). London: Psychology Press.

García-Madruga, J. A, Gutiérrez, F, Carriedo, N., Vila, J. O. & Luzón J.M. (2007). Mental Models in Propositional Reasoning and Working Memory's Central Executive. *Thinking and Reasoning* (in press)

García-Madruga, J. A., Gutiérrez, F., Carriedo, N., Moreno, S. & Johnson-Laird, P. N. (2002). Mental models in deductive reasoning. *The Spanish Journal of Psychology, 5*, 125-140.

Geis, M.L. (1973). If and unless. In B.B. Kachu, R.B. Lees, Y. Malkiel, A. Pietrangeli, & S. Saporta (Eds). *Issues in linguistics* (pp. 231-253). Urbana Chicago: University of Illinois Press.

Grosset, N. and Barrouillet, P. (2003). On the nature of mental models of conditional: The case of *If*, *If then* and *Only if*. *Thinking and Reasoning, 9*, 289-306.

Holyoak, K.J., & Cheng, P. (1995). Pragmatic reasoning with a point of view: A response. *Thinking and Reasoning, 1*, 289-313, 373-388.

Jackson, F. (1987). *Conditionals*. Oxford: Blackwell.

Johnson-Laird, P. N. & Byrne, R. M. J. (1989). Only reasoning. *Journal of Memory and Language. 28*, 313-330.

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

Johnson-Laird, P. N. & Byrne, R. M. J. (2002). Conditionals: a theory of meaning, inference, and pragmatics. *Psychological Review.* 109, 646-678.

Johnson-Laird, P.N., Byrne, R., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review, 99*, 418-439.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In Kahneman, D. P. Slovic, and Tversky, A. (Eds.). *Judgment under uncertainty: heuristics and biases.* pp. 201-208. New York: Cambridge University Press.

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

Oberauer, K., Hörnig, Weidenfeld, A., and Wilhelm, O. (2005). Effects of directionality in deductive reasoning: II. Premise integration and conclusion evaluation. *Quarterly Journal of Experimental Psychology, 58A*, 1225-1247.

Oberauer, K. and Wilhelm, O. (2000). Effects of directionality in deductive reasoning: I. The comprehension premises. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 1702-1712.

Pollack, J. L. (1986). *Subjunctive reasoning*. Dordrecht, The Netherlands: Reidel.

Quine, W.V. (1972). *Methods of logic* (third edition). New York: Holt, Rinehart & Winston.

Reichenback, H. (1947). *Elements of symbolic logic*. New York: Free Press.

Revlin, R., Cate, C. & Rouss, T. (2001). Reasoning counterfactually: Combining and rending. *Memory and Cognition, 2*, 1196-1208.

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

Roberge, J.J. (1978). Linguistic and psychometric factors in propositional reasoning. *Quarterly Journal of Experimental Psychology,* 30, 705-716.

Santamaría, C. & Espino, O. (2002). Conditionals and directionality: on the meaning of 'if' vs. 'only if'. *Quarterly Journal of Experimental Psychology, 55A* (1), 41-57.

Santamaría, C., Espino, O. and Byrne, R.M.J. (2005). Counterfactual and semifactual conditionals prime alternative possibilities. *Journal of Experimental Psychology*: *Learning, Memory and Cognition.* *31*, 1149 – 1154.

Schaeken, W., García-Madruga, J. A. & d'Ydewalle (1997). *Unless reasoning*. Unpublished Paper.

Stalnaker, R.C. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory*. Oxford: Basil Blackwell.

Thompson, V. & Byrne, R.M.J. (2002). Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory & Cognition.* 28, 1154-1170.

von Fintel, K. (1991). Exceptive conditionals: the meaning of *unless. North Eastern Linguistics Society, 22*, 135-48.

Wright, P. & Hull, A.J. (1986). Answering questions about negative conditionals. *Journal of Memory and Language,* 25, 691-709.

Wright, P. & Hull, A.J. (1988). Reading to do: creating contingent action plans. *British Journal of Psychology ,* 79, 187-211.

APPENDIX

Examples of factual and counterfactual conditionals used in the experiments in the form 'if not', unless and 'only if', in Portuguese, Spanish and English.

Factual
If not
Portuguese: Se a Cristina não está em Lisboa, o Sergio não está em Granada
Spanish: Si Cristina no está en Lisboa, Sergio no está en Granada
English: If Cristina is not in Lisboa, Sergio is not in Granada

Unless
Portuguese: O Sergio não está em Granada, a não ser que a Cristina esteja em Lisboa
Spanish: Sergio no está en Granada a menos que Cristina esté en Lisboa
English: Sergio is not in Granada unless Cristina is in Lisboa

Only if
Portuguese: O Sergio está em Granada, apenas se a Cristina estiver no Lisboa
Spanish: Sergio está en Granada sólo si Cristina está en Lisboa
English: Sergio is in Granada only if Cristina is in Lisboa

Counterfactual
If not
Portuguese: Se a Cristina não tivesse estado em Lisboa, o Sergio não teria estado em Granada
Spanish: Si Cristina no hubiese estado en Lisboa, Sergio no habría estado en Granada
English: If Cristina had not been in Lisboa, Sergio would have not been in Granada

Unless
Portuguese: O Sergio não teria estado em Granada, a não ser que a Cristina tivesse estado em Lisboa
Spanish: Sergio no habría estado en Granada a menos que Cristina hubiese estado en Lisboa
English: Sergio would had not been in Granada unless Cristina had not been in Lisboa

Only if
Portuguese: O Sergio teria estado em Granada, apenas se a Cristina tivesse estado em Lisboa
Spanish: Sergio habría estado en Granada sólo si Cristina hubiera estado en Lisboa
English: Sergio would have been in Granada only if Cristina had been in Lisboa

(Manuscript received: 19 December 2007; accepted: 11 June 2008)