Reasoning with Exceptive Conditionals: 
The case of ‘Except If’

O. Espino*, I. Sánchez-Curbelo and G.M. Ramírez

Universidad de La Laguna, Tenerife, Spain.

In this paper we outline a mental model theory (Johnson-Laird & Byrne, 1991, 2002) of reasoning about the conditional except if. We report two experiments showing that the exceptive conditional except if exerted certain forms of semantic modulation and determined the inferences that individuals draw in an inference task (Experiment 1) and in a truth table task (Experiment 2). In Experiment 1 we found that there were no reliable differences between the percentage of modus ponens and modus tollens with the conditional except if but that the differences were reliable with the conditional if not, then. In Experiment 2, participants selected the possibilities ‘A & not-B’ and ‘not-A & B’ more frequently than the possibilities ‘A & B’ and ‘not-A & not-B’ with the conditional ‘B except if A’, but they selected the possibility ‘not-A & B’ more frequently than other possibilities (‘A & not-B’, ‘A & B’, ‘not-A & not-B’) with the conditionals ‘B if not-A’ and ‘if not-A, B’. The implications of these results are discussed in the context of recent psychological and linguistic theories in respect of the meaning of except if.

In this paper, we will focus on how people understand and think with the exceptive conditional, found in constructions such as ‘the women can take this drug except if she is allergic to the penicillin’. Normally, people use this type of conditional when they want to express when the conditions under the consequent (the women can take this drug) will not be carried out (she is allergic to the penicillin). Although the mental representations underlying comprehension and reasoning of the negative exceptive

* This research was supported by the Research Project PSI2011-26386 from the Spanish Ministry of Science and Innovation (Ministerio de Ciencia e Innovación). Contact information: Orlando Espino. Departamento de Psicología Cognitiva, Social y Organizacional. Facultad de Psicología. Universidad de la Laguna. Campus de Guajara, 38205. Tenerife. Spain. Phone 34 922317505. Fax 34 922317461. E-mail: oespinom@ull.es
conditional *unless* have attracted the attention of linguists (Dancygier & Sweetser, 2005; Declerk & Reed, 2000; Lycan, 2001), philosophers (Fillenbaum, 1986; Quine, 1972; Reichenbach, 1947) and cognitive psychologists (Carriedo, García-Madruga, Moreno-Ríos, & Gutiérrez, 1999; García-Madruga, Carriedo, Moreno-Ríos, Gutiérrez, & Schacken, 2008; García-Madruga, Carriedo, & Moreno-Ríos, 2011; García-Madruga, Gutiérrez, Carriedo, Moreno-Ríos, & Johnson-Laird, 2002; García-Madruga, Moreno-Ríos, Quelhas, & Juhos, 2009; Gómez-Veiga, García-Madruga, & Moreno-Ríos, 2012; Wright & Hull, 1986), interest in the study of the exceptive conditional *except if* has been minimal (Espino, Sánchez-Curbelo, García, & Estupiñan, 2013).

There are different reasons why the exceptive conditional *except if* has attracted little attention from psycholinguists and cognitive psychologists. The first reason has to do with the idea that for some authors, this type of conditional is logically equivalent to *if* clauses in the sense that traditionally it is assumed that the exceptive conditional ‘B except if A’ is equivalent to the conditional ‘B, if not-A’ (or ‘if not-A, then B’) (Montolío, 2000). However, in this paper we claim that the exceptive conditional ‘B except if A’ is not logically equivalent to the conditional ‘B, if not-A’ for a number of reasons. First, exceptive conditionals establish a much more precise and restricted relationship between the main clause (‘B’) and the subordinate clause (‘A’) than the conditional *if not, then*. In other words, *except if* is more specific in meaning and less abstract than the conditional *if* and its negative counterpart *if not, then*. This specific relationship between the main clause (‘B’) and the subordinate clause (‘A’) in exceptive conditionals is so strong that it cannot be cancelled without leading to a semantically and pragmatically unacceptable utterance (Montolío, 2000).

Montolío (1999, 2000) has claimed that this behavior is similar to the connective *if and only if* and consequently exceptive conditionals are better understood as biconditionals. The second reason why the exceptive conditional ‘B except if A’ is not equivalent to the conditional ‘B, if not-A’ has to do with the fact that *if not, then* clauses can be iterated in coordinate structures while *except if* clauses cannot. For example, the following coordinate sentences ‘I will go out with you tonight if you do not drink and if you do not drive very fast’ are acceptable, but when we use the exceptive conditional ‘I will go out with you tonight except if you drink and except if you drive very fast’, they are not. Montolío (2000) pointed out that the reason behind the unacceptability of iterated exceptive conditionals lies in the fact that it is not possible to present the only circumstance under which something will not happen and then add another exceptional condition.
Reasoning with Exceptive Conditionals

One important practical implication for the use of *except if* as opposed to *if not, then* is pragmatic. Generally, the semantics of the ‘exceptive conditional’, such as *except if* and *unless*, make it appropriate for it to appear in deterrent contexts while at the same time making it unnatural-sounding in some other contexts (for example, the causal context) in which an *if* clause can be used. For example, the following sentences sound natural with an *if* clause:

If the vase drops, then it will break
But unnatural with *except if*:
The vase will break except if the vase is not dropped.

Also, exceptive sentences sound unnatural in inducement contexts, as the following example shows:

I will give you 10 euro except if you are not quiet
However, this sentences sound natural with *if-clause*:
If you are quiet, I will give you 10 euro.

Another important aspect that distinguishes the exceptive conditional from the indicative conditional has to do with the fact that people find it more difficult to make inferences from exceptive conditionals than from indicative conditionals. For example, Schaeken, García-Madruga, & d’Ydewalle (1997) found that solving conditional inferences with *unless* was harder than with *only if*. The special difficulty of *unless* conditionals results mainly from their tendency to produce asymmetric conclusions, while this is not the case for *only if*. As an example of these kinds of asymmetric responses, some people gave the incorrect response ‘B’ to the following set of premises: ‘not-A unless B’, and ‘not-A’. Schaeken et al. (1997) have claimed that a hypothetical explanation could be that these reasoners were using a shortcut strategy that consists of matching the two terms that appear in the statement. Also, García-Madruga et al. (2008) found that there was a relevant percentage of asymmetric conclusions with *unless* conditionals, reliably higher than those found with the other conditional formulations (*if, then; if not, then; and only if*). Montolío (2000) claims that the exceptive conditionals ‘imply a greater cognitive-processing complexity in that to the uncertainty of the truth value they add a negative relationship between the two clauses’ (page 163). In light of these facts, it is not surprising that some authors have claimed that the exceptive conditional *except if* is not equivalent to the conditional *if not, then* (Dancygier, 2002; Dancygier & Sweetser, 2005). However, these authors have not provided experimental evidence to support their claims. One of our goals in this
paper will be to offer experimental evidence to show that the two types of conditional are not semantically equivalent.

The second reason why exceptive conditionals, such as except if, have attracted little attention from psycholinguists and cognitive psychologists has to do with the idea that for some authors, except if and unless are semantically equivalent (Declerck & Reed, 2000; Geis, 1973; Montolío, 2000). For instance, Montolío (2000) has claimed that the conditional connective ‘excepto si’ (except if) is semantically analogous to the connective ‘a menos que’ (unless). As a consequence of this assumption, there has been a preference to study exceptive conditionals using the utterance unless more than except if. Contrary to these authors, we suggest that the exceptive conditionals except if and unless could mean different things. In the Spanish language, the conditional unless requires the subjunctive mood in the subordinate clause, while the conditional except if can only be combined with the indicative. In both English and Spanish, the indicative mood is used to express factual information, certainty, and objectivity while the subjunctive mood conveys wishes, conjectures and uncertainties. We suggest that the exceptive conditional except if is more precise than the exceptive conditional unless in expressing the exceptive circumstance. This idea is close to the views expressed by Dancygier & Sweetser (2005) when they argue that except if and unless are different in some circumstances, because they in fact establish exceptive spaces of different kinds, and to Dancygier’s (2002) affirmation that except if can only loosely be assumed to be synonymous with unless. In brief, in this paper we are assuming that except if has a biconditional meaning and that it describes exceptive possibility more precisely than the conditional unless.

**Exceptive conditional and mental model theory**

The main goal in this paper will be to outline a mental model theory (Johnson-Laird & Byrne, 1991, 2002) of reasoning about the conditional except if. Mental model theory – or model theory for short – claims that reasoning depends not on logical form but on mental models, which are psychological analogues of the models that represent content in logic. According to this theory, individuals use the meaning of words, the grammatical structure of sentences and their knowledge to construct models of the possibilities to which propositions refer, and a conclusion is considered to be valid if it holds in all these models (Johnson-Laird & Byrne, 2002; Johnson-Laird, 2006). In respect of the model theory view, there are several key principles that govern the mental representations that people construct. The first principle is that people keep in mind only true
possibilities (‘truth principle’). For instance, people may understand ‘if there is a circle, then there is a triangle’ by thinking about the true possibilities, ‘there is a circle and there is a triangle’, ‘there is not a circle and there is a triangle’ and ‘there is not a circle and there is not a triangle’ but not the false possibility ‘there is a circle and there is not a triangle’ (Johnson-Laird, 2006; Johnson-Laird & Byrne, 2002). The second principle claims that people keep in mind few true possibilities (‘parsimony principle’) because of the constraints of working memory (Johnson-Laird, Byrne, & Schaeken, 1992). Hence, when people reason from a basic conditional, such as ‘if there is a circle, then there is a triangle’, they normally construct a single mental model that represents the first possibility above in which the conditional’s antecedent (circle) and its consequent (triangle) are both true and an implicit mental model (as shown by the ellipsis) that represents the other possibilities in which the antecedent is false:

\[
\text{circle} \quad \text{triangle} \quad \ldots
\]

If it is required, people can ‘flesh out’ their understanding of the conditional in order to think about the other possibilities and make them fully explicit. A third principle claims that for some conditionals people are required to think about two possibilities (‘dual possibilities principle’). For instance, Byrne (2005, 2007) and Johnson-Laird & Byrne (1989) claim that the assertion ‘there is a circle only if there is a triangle’ prompts people to think of dual possibilities (circle & triangle; not-circle & not-triangle). As a result, they can readily make both MP and MT inferences. The fourth principle claims that the interpretation of the conditional is subject to a process of semantic and pragmatic modulation. It is claimed that the meaning of the clauses in conditionals and co-referential relations between them can modulate the core meaning in a process of semantic modulation and the knowledge about the context and the topic of the conditional can modulate the core meaning in a process of pragmatic modulation (Johnson-Laird & Byrne, 2002; Quelhas, Johnson-Laird & Juhos, 2010). In this paper, we claim that the interpretation of a conditional can be influenced by the type of linguistic expression (such as, except if, on condition that, unless, etc.) employed to express the conditional and our goal will be to show that the exceptive expression except if can exert certain forms of modulation and determine the inferences that individuals draw in a truth table task and in an inference task. Our intention is to show that the
expression ‘B except if A’ (or ‘except if A, B’), which is logically equivalent to ‘B, if not-A’ (or ‘if not-A then B’), can block the construction of the true possibility ‘A and B’. This hypothesis is based on the idea that the event described in exceptive clauses such as B except if A should be understood to be an exceptional circumstance (‘A’) under which the situation described in the main clause (‘B’) will not occur (Dancygier & Sweetser, 2005). We also claim that when people understand an exceptive conditional, such as ‘B except if A’ (or ‘except if A, B’), they have in mind just two models: ‘A and not-B’ and ‘not-A and B’. Our proposal is close to this author’s claim that the exceptive conditional has a biconditional meaning (Montolío, 2000). In keeping with this idea, we predict that in a truth table task participants will tend to accept the possibility ‘not-A & B’ as frequently as the possibility ‘A & not-B’ with the exceptive conditional ‘B except if A’ (or ‘except if A, B’). On the other hand, we predict that in the conditional ‘B, if not-A’ and ‘if not-A, then B’, participants will tend to accept the possibility ‘not-A & B’ more frequently than the possibility ‘A & not-B’, because the first possibility is part of the initial representation but not the second possibility. In a truth table task, participants have to evaluate if different possibilities (for example, ‘not-A and B’) are ‘true’, ‘false’ or ‘it is not possible to know’ for the conditional rule (for example, ‘if not-A, then B’). Table 1 shows the initial true possibilities and the implicit possibilities that people should keep in mind for the conditional if not, then and except if.

In an inference task participants have to construct or evaluate conclusions from the application of four conditional rules: modus ponens (MP), denial of antecedent (DA), affirmation of the consequent (AC), and modus tollens (MT) (Table 2). We predict that in an inference task there will be differences in accuracy between MP and MT inferences for if not, then statements, since modus tollens cannot be drawn from the initial model. On the other hand, we predict no differences between MP and MT inferences for except if, since both inferences can be drawn from the initial representation. Also, we predict that people should tend to accept more frequently the affirmation of consequent (AC) than the denial of antecedent (DA) inference in the conditional if not, then, due to the fact that AC can be drawn from the initial model but DA could not be obtained from the initial model. On the other hand, it is predicted that participants will tend to accept the AC inferences and DA inferences equally frequently in the exceptive conditional except if. In the conditional except if, both inferences could be obtained from the initial models.
Table 1. Proposed possibilities for the conditional if not, then (‘B, if not-A’ or ‘if not-A, then B’) and for the exceptive conditional except if (‘B except if A’ or ‘except if A, B’).

<table>
<thead>
<tr>
<th>Conditional Form</th>
<th>Initial possibilities</th>
<th>Full explicit possibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>B except if A</td>
<td>B &amp; not-A</td>
<td>B &amp; not-A</td>
</tr>
<tr>
<td></td>
<td>not-B &amp; A</td>
<td>not-B &amp; A</td>
</tr>
<tr>
<td>B, if not-A</td>
<td>B &amp; not-A</td>
<td>B &amp; not-A</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>not-B &amp; A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B &amp; A</td>
</tr>
<tr>
<td>except if A, B</td>
<td>A &amp; not-B</td>
<td>A &amp; not-B</td>
</tr>
<tr>
<td></td>
<td>not-A &amp; B</td>
<td>not-A &amp; B</td>
</tr>
<tr>
<td>if not-A, B</td>
<td>not-A &amp; B</td>
<td>A &amp; not-B</td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>A &amp; B</td>
</tr>
</tbody>
</table>

Note. Each horizontal row denotes a model of a separate possibility.

**Alternative theoretical view**

In this paper we contrast our predictions against alternative theories of reasoning, such as formal rule theories and hypothetical thinking. Formal rule theories (Braine & O’Brien, 1998; Rips, 1994) postulate that our mind contains a set of formal rules of inference akin to a logical calculus. The rules apply once the logical form of propositions has been matched to them. In other words, reasoning proceeds like a proof in formal logic, by application of standard inferential rules or schemas. One of the most important predictions of the formal rule theories is that if a problem can be solved with a simple rule, it will be easier than if a complex rule is needed. According to this theory, people should make more MP inferences than MT inferences because MP is a simple rule of inference while MT is a complex rule (reductio ad absurdum). This rule postulates that if a premise is supposed and then a contradiction inferred, the negation of the supposition is a valid conclusion.
Table 2. The four inferences for the four linguistic forms.

<table>
<thead>
<tr>
<th>Inferences</th>
<th>If not-A, B</th>
<th>B, if not-A</th>
<th>B, except if A</th>
<th>Except if A, B</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP</td>
<td></td>
<td></td>
<td>not-A/ therefore, B</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td></td>
<td></td>
<td>B/ therefore, not-A</td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td></td>
<td></td>
<td>A/ therefore, not-B</td>
<td></td>
</tr>
<tr>
<td>MT</td>
<td></td>
<td></td>
<td>not-B/ therefore, A</td>
<td></td>
</tr>
</tbody>
</table>

Note. MP = modus ponens, AC = affirmation of the consequent, MT = modus tollens, DA = denial of antecedent.

On the other hand, Suppositional theory (Evans & Over, 2004; Evans, Over, & Handley, 2005) has claimed that the comprehension of conditionals leads to the construction of a single mental model or possibility (singularity principle). In the exceptive conditional ‘B except if A’, the single possibility is relative to the true antecedent (not-A) but not the false antecedent (A). Then, Suppositional theory should predict that participants should tend to accept one possibility (whose antecedent is true) more than the other possibility (whose antecedent is false). A second claim of Suppositional theory is that people make the MT inferences by the rule ‘*reductio ad absurdum*’. According to this theory, it is predicted that people should make more MP inferences than MT inferences in factual conditional and exceptive conditional contexts.

We will employ an inference task and truth table task to infer what kind of mental representation people have in mind when they understand and think with the exceptive conditional *except if*. 
EXPERIMENT 1

The objective of this experiment was to compare people’s reasoning with the four logically equivalent conditional formulations: ‘If not-A then B’, ‘B if not-A’, ‘B except if A’, and ‘except if A, B’. As we analyzed above, people build only one complete initial possibility or model for if not, then, whereas they construct two complete possibilities for except if (see Table 1). From these assumptions we predict that there will be differences in accuracy between MP and MT inferences for if not, then statements, since MT cannot be drawn from the initial model. On the other hand, we predict no differences between MP and MT inferences for except if, since both inferences can be drawn from the initial representation. It is predicted that people should tend to accept more frequently AC than DA inferences in the conditional if not, then, due to the fact that AC can be obtained from the initial model but not DA. Also, we predict participants will tend to accept the AC and DA inferences equally frequently in the exceptive conditional except if. In the conditional except if, both inferences could be obtained from the initial models. Finally, given the special difficulty of exceptive conditional statements examined in prior studies, we predict that reasoners will tend to give more asymmetric conclusions with except if than with the conditional if not, then.

On the other hand, formal rule theories (Braine & O’Brien, 1998; Rips, 1994) and Suppositional theory (Evans & Over, 2004; Evans et al., 2005) predict that people should make more MP inferences than MT inferences in both types of conditional (except if and if not, then), because MP inference is a simple rule of inference while MT inference is a complex rule.

Using the exceptive conditional with two expressions (‘B except if A’ and ‘except if A, B’) not only allows us to check if the exceptive conditional is semantically equivalent to the indicative conditional (such as, ‘B if not-A’ and ‘if not-A then B’) but also to check if the exceptive conditionals ‘except if A, B’ and ‘B, except if A’ are semantically equal (Declerck & Reed, 2000).

METHOD

Participants. The 92 participants took part in the experiment, 46 in each group. The participants were undergraduate students at the University of La Laguna, Tenerife, Spain. None of the participants in this or any of the subsequent experiments reported having received formal training in logic.
**Design.** A 2x2x4 mixed-subject design was used in this experiment. The between-subject variable was type of conditional: one group received ‘if not’ problems and the other group received ‘except if’ problems. In the within-subject design, the first independent variable was locus of the connective, with two levels: connective at the beginning (e.g., ‘if not-A then B’ and ‘except if A, B’) and connective in the middle (e.g., ‘B, if not-A’, and ‘B except if A’). The second independent variable in the within-subject design was type of inference, with four levels: *modus ponens*, denial of antecedent, affirmation of consequent, *modus tollens*. The dependent variable was the percentage of inference accepted as valid.

**Materials and Procedure.** We tested the participants in groups. We gave each participant a booklet, the first page of which contained instructions explaining the task with reference to the disjunction argument. We gave the participants the following instructions:

«This task is designed to test your understanding of logical rules. On the following pages you will be presented with a series of problems. In each problem a rule will be presented followed by a conclusion. For each problem you must indicate whether or not the conclusion necessarily follows given the rule that precedes it, or whether it is not possible to know. A conclusion is necessarily true when the conclusion must follow given the truth of the rule. »

They were asked to read each problem carefully and to work from beginning to end at their own pace without changing any response or skipping any items. One group received eight problems with ‘if not’ and the other group received eight problems with ‘except if’. They received the problems in different random order, one problem for each experimental condition. The content used in this experiment was about geometrical figures, such as ‘on the table there is a circle except if there is a triangle’. As an example of a problem is as follows:

On the table there is a circle except if there is a triangle
On the table there is a circle
Does it follow that:

There is a triangle
There is not a triangle
It is not possible to know
Participants were asked to indicate their response by circling either 'yes', 'no' or 'it is not possible to know' for each problem.

RESULTS AND DISCUSSION

Table 3 presents the percentage endorsements of the inferences as a function of the type of conditional (‘B, if not-A’, ‘if not-A then B’, ‘B except if A’ and ‘except if A, B’) and type of inference (MP, DA, AC and MT). We performed a 2 (type of conditional: ‘if not’ versus ‘except if’) x 2 (locus of the connective: at the beginning versus in the middle) x 4 (type of inference: MP, DA, AC and MT) analysis of variance (ANOVA) with repeated measures in the last two factors, and the first factor as a between-subject variable. The Greenhouse–Geisser correction for the violation of the sphericity assumption was used in this experiment and for each analysis presented in this paper. There was interaction between type of conditional and locus of the connective, $F(1, 90) = 6.88$, $MSE = .23$, $p < .015$, $\eta^2 = .07$, and between locus of the connective and type of inference, $F(2.650, 238.4) = 10.37$, $MSE = .14$, $p < .001$, $\eta^2 = .10$. The triple interaction was reliable, $F(3, 270) = 9.69$, $MSE = .15$, $p < .001$, $\eta^2 = .10$. There was a main effect of locus of the connective, $F(1, 90) = 14.88$, $MSE = .23$, $p < .001$, $\eta^2 = .14$, and of type of inference, $F(3, 270) = 3.85$, $MSE = .13$, $p < .015$, $\eta^2 = .04$. The interaction between type of conditional and type of inference was not reliable, $F(3, 270) = 2.49$, $MSE = .13$, $p = .06$, $\eta^2 = .03$. Finally, there was main effect for the type of conditional, $F(1, 90) = 5.55$, $MSE = .56$, $p < .025$, $\eta^2 = .06$.

To test the prediction that the percentages of MP inferences would be higher than MT inferences in ‘B, if not-A’ and ‘if not-A then B’ but not in ‘B except if A’ and ‘except if A, B’, a series of planned comparisons was carried out. As we predicted, there was a reliable difference between MP and MT inferences in ‘B, if not-A’ (87% vs 54%; $t(45) = 3.69$; Bonferroni, $p < .05$) and in ‘if not-A, then B’ (91% vs 50%; $t(54) = 5.18$; Bonferroni, $p < .05$), and there were no reliable differences between MP and MT inferences in ‘B except if A’ (76% vs 63%; $t(45) = 2.00$, Bonferroni $p > .05$) or in ‘except if A, B’ (46% vs 57%; $t(45) = 1.52$, Bonferroni $p > .05$). As we predicted, no differences were found between AC and DA inferences on ‘B except if A’ (76% vs 83%; $t(45) = 1.13$, Bonferroni $p > .05$), ‘except if A, B’ (52% vs 52%; $t(45) = 0$, Bonferroni $p > .05$). As we also predicted, there were reliable differences between AC and DA inferences for the conditional ‘B, if not-A’ (96% vs 76%; $t(45) = .330$, Bonferroni $p > .05$) but not for the conditional ‘if not-A, then B’ (83% vs 72%; $t(45) = 1.30$,
Bonferroni $p > .05$), although the differences were in the predicted direction.

Table 3. Percentages of endorsements of the inferences as a function of the type of conditional (‘B if not-A’, ‘If not-A then B’, ‘B except if A’, ‘Except if A, B’) and type of inference (modus ponens (not-A; therefore B), denial of antecedent (A; therefore not-B), affirmation of consequent (B; therefore not-A), modus tollens (not-B, therefore A)) in Experiment 1. Asymmetric responses are shown in brackets.

<table>
<thead>
<tr>
<th></th>
<th>B, if not-A</th>
<th>If not-A, then B</th>
<th>B, except if A</th>
<th>Except if A, B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modus Ponens</td>
<td>87 (2)</td>
<td>91 (5)</td>
<td>76 (11)</td>
<td>46 (39)</td>
</tr>
<tr>
<td>Denial Antecedent</td>
<td>76 (2)</td>
<td>72 (4)</td>
<td>83 (11)</td>
<td>52 (45)</td>
</tr>
<tr>
<td>Affirmation Consequent</td>
<td>96 (0)</td>
<td>83 (2)</td>
<td>76 (15)</td>
<td>52 (28)</td>
</tr>
<tr>
<td>Modus Tollens</td>
<td>54 (7)</td>
<td>50 (9)</td>
<td>63 (17)</td>
<td>57 (35)</td>
</tr>
<tr>
<td>Total</td>
<td>78 (3)</td>
<td>74 (5)</td>
<td>75 (14)</td>
<td>52 (37)</td>
</tr>
</tbody>
</table>

As we predicted, there were no reliable differences between MP and MT inferences and between AC and DA inferences for exceptive conditionals. The absence of differences between MP and MT and between AC and DA inferences on the conditional except if suggests that participants reason with ‘B except if A’ and ‘except if A, B’ by initially envisaging two possibilities: ‘not-A & B’ and ‘A & not-B’. On the other hand, as we predicted, there were reliable differences between MP and MT on the conditional if not, then. This data suggest that MP is part of the initial representation while MT is not part of the initial representation.

As we predicted, there were reliable differences between AC and DA inferences for the conditional ‘B, if not A’, but there were no reliable
Reasoning with Exceptive Conditionals

There are differences between AC and DA inferences for the conditional ‘if not A, B’, although the differences were in the predicted direction. We explain the absence of difference between AC and DA by claiming that the advantage of being part of the initial model can be counteracted by the backward processing. According to Mental Model theory (Johnson-Laird & Byrne, 1991) forward and backward inferences depend on the way the information enters working memory. For example, for the major premise ‘if not-A, B’, the information about ‘not-A’ enters working memory before the information about ‘B’. Some research has shown that in the conditional rule ‘if A, then B’, participants make more forward inferences (from A to B) than backward inferences (from B to A; Evans, 1977, 1993; Evans & Beck, 1981). As well, Barrouillet, Grosset, and Lecas (2000, Exp. 1 and 2) found that participants took longer to endorse inferences from the implicit model than from the initial model. In the conditional ‘if not-A, B’, ‘B’ is part of the initial model but it is processed in a backward direction when participants have to make the AC inference (if not-A, B; B; therefore not-A). The advantage of AC of being part of the initial model is counteracted because it is processed in backward direction. On the other hand, ‘A’ is not part of the initial model for the conditional ‘if not-A, B’ but it is processed in forward direction if participants make the DA inference (if not-A, B; A; therefore not-B). The advantage of DA of being processed in a forward direction is counteracted because it is not part of the initial models. In the conditional ‘B if not-A’, the AC inference can be obtained from the initial models and it is processed in forward direction while the DA inference cannot be obtained from the initial model and it has to be processed in backward direction.

The fact that most participants accept MP as frequently as MT inferences in ‘B except if A’ corroborates the model theory, which postulates that both inferences can be obtained from the initial representation. However, it presents a difficulty to the theories based on formal rules of inference (Braine & O’Brien, 1998; Rips, 1994), which contain no rules to explain the absence of differences between MP and MT. Also, this result cannot be explained by Suppositional theory (Evans & Over, 2004; Evans et al., 2005), which predicts that MP inferences should be accepted more frequently than MT inferences.

We performed a 2 (type of conditional: ‘if not’ versus ‘except if’) x 2 (locus of the connective: at the beginning versus in the middle) x 4 (type of inference: MP, DA, AC and MT) analysis of variance (ANOVA) with repeated measures in the last two factors for asymmetric responses and the first factor as between-subject variable. There are four asymmetric responses for the four inferences: one is to accept the inferences ‘not-B’
from the categorical premise ‘not-A’; the second is to accept the inferences ‘not-A’ from the categorical premise ‘not-B’; the third is to accept the inferences ‘B’ from the categorical premise ‘A’; and the fourth is to accept the inferences ‘A’ from the categorical premise ‘B’. There was a main effect of type of conditional, $F(1, 90) = 27.91$, $MSE = .30$, $p < .001$, $\eta^2 = .34$ in the sense that participants gave more asymmetric responses in exceptive than in if conditional (25% versus 4%). There was interaction between type of conditional and locus of the connective, $F(1, 90) = 15.75$, $MSE = .13$, $p < .001$, $\eta^2 = .15$, in the sense that participants gave more asymmetric responses in the conditional ‘except if A, B’ than in the conditional ‘B except if A’ (37% versus 14%; $t(45) = 4.61$, $p < .001$) but they gave similar percentages of asymmetric responses in the conditional ‘B, if not-A’ as in the conditional ‘if not-A, then B’ (3% versus 5%; $t(45) = 1.27$, $p = .21$). There was a main effect of locus of connective, $F(1, 90) = 22.87$, $MSE = .13$, $p < .001$, $\eta^2 = .20$. The interaction was not reliable between type of inference and type of connective, $F(3, 270) = 1.73$, $MSE = .07$, $p = .16$, $\eta^2 = .02$, or between type of inference and locus of the connective, $F(3, 270) = 1.19$, $MSE = .06$, $p = .31$, $\eta^2 = .01$. There was no main effect for type of inference, $F(3, 270) = 1.81$, $MSE = .07$, $p = .13$, $\eta^2 = .02$. The triple interaction was not reliable ($F > 1$).

The high presence of asymmetric responses in the conditional ‘except if A, B’ shows that the participants had difficulty understanding and thinking with this connective. We suggest that participants could not understand the meaning of this connective because in the Spanish language the clauses with except if most commonly follow the main clauses (‘B except if A’) and conditionals with pre-posed clauses (‘except if A, B’) are very rare and unusual. We suggest that when participants have to think with an unusual expression they could use a shortcut strategy that consists of matching the two terms that appear in the statement ‘except if A, B’. Finally, these results run counter to claims that there is no semantic difference between ‘except if A, B’ and ‘B except if A’ (Declerck & Reed, 2000).

**EXPERIMENT 2**

The objective of this experiment was to compare people’s reasoning with the four logically equivalent conditional formulations: ‘if not-A then B’, ‘B if not-A’, ‘B except if A’, and ‘except if A, B’. Our main assumption, based on Mental Model Theory (Johnson-Laird & Byrne, 2002), is that with the conditional ‘B if not-A’, people build one initial
model (‘not-A & B’) whereas they construct two initial models or possibilities for the exceptive conditional ‘B except if A’ (‘not-A & B’ and ‘not-B & A’). From these assumptions we predict that participants will tend to accept the possibility ‘not-A & B’ as frequently as the possibility ‘A & not-B’ for the conditional except if. On the other hand, we predict that in the conditionals ‘B, if not-A’ and ‘if not-A, then B’, participants will tend to accept the possibility ‘not-A & B’ more frequently than the possibility ‘A & not-B’. This last prediction is based on the idea that the first possibility is part of the initial representation while the second possibility is not part of initial representation. To test these predictions, we examined the inferences that people made with ‘B except if A’ and ‘except if A, B’, as compared with ‘B, if not-A’ and ‘if not-A, then B’ in a truth table task.

We contrast our predictions against Suppositional theory (Evans & Over, 2004; Evans et al., 2005). Suppositional theory predicts that, in the exceptive conditional ‘B except if A’ and ‘except if A, B’, people only think about the possibilities that include the true antecedent (not-A) but they do not think about the possibilities that include the false antecedent (A). Consequently, Suppositional theory should predict that participants should tend to accept, in a truth table task, the option in which the antecedent is true ‘not-A & B’ more frequently than an option in which the antecedent is false ‘A & not-B’.

**METHOD**

Participants. The 41 participants who took part in the experiment were undergraduate students at the University of La Laguna, Tenerife, Spain.

Design. A 4x4 within subject design was used in this experiment. The first independent variable was type of conditional with four levels: ‘B, if not-A’, ‘If not A, then B’, ‘B except if A’ and ‘except if A, B’. The second independent variable was type of conjunction with four levels: ‘A & B’, ‘A & not-B’, ‘not-A & B’ and ‘not-A & not-B’. The dependent variable was the percentage of responses accepted as valid.

Materials and Procedures. Participants received a booklet consisting of 5 pages. The first page contained the following instructions:
This task is designed to test your understanding of logical rules. On the following pages you will be presented with a series of problems. In each problem, a rule will be presented followed by a series of outcomes. For each problem you must indicate whether each outcome is true, false or not possible to know given the truth of the rule.

On this page they were given an example with a biconditional. On each of the following pages, participants received four different types of conditionals in random fashion. The content used in this experiment was about geometrical figures, such as ‘on the table there is a circle except if there is a triangle’. Following the presentation of the rule, participants were required to indicate whether each of the following four truth table cases was true, false or not possible to know given the truth of the rule:

- ‘There is a circle and there is a triangle’.
- ‘There is a circle and there is not a triangle’.
- ‘There is not a circle and there is a triangle’.
- ‘There is not a circle and there is not a triangle’.

They indicated their response by circling either ‘true’, ‘false’ or ‘it is not possible to know’. They received 16 problems in a different random order, one problem for each experimental condition.

RESULTS AND DISCUSSION

Table 4 presents the percentage of participants who indicated that each truth table case was true with respect to the rule for the different conditionals. We carried out a 4 (type of conditional: ‘B, if not-A’, ‘if not A, then B’, ‘B except if A’ and ‘except if A, B’) x 4 (type of conjunction: ‘A & B’, ‘A & not-B’, ‘not-A & B’, ‘not-A & not-B’) analysis of variance (ANOVA) with repeated measures on both factors. There was an interaction between type of conditional and type of conjunction, $F(5.438, 217.528) = 18.06, MSE = .23, p < .001, \eta^2 = .31$. As we predicted, participants accepted the conjunction ‘not-A & B’ more frequently than the conjunction ‘A & not-B’ (93% vs 46%, t (40) = 5.87; Bonferroni, $p < .005$) in the conditional ‘B if not-A’. Also, it was found that they accepted the conjunction ‘not-A & B’ more frequently than the conjunctions ‘A & B’ and ‘not-A & not-B’ (93% vs 5%, t (40) = 14.05; 93% vs 15%, t(40) = 11.92; Bonferroni, $p < .005$).
Moreover, as we predicted, they accepted the conjunction ‘not-A & B’ more than the conjunction ‘A & not-B’ in the conditional ‘if not-A, then B’ (93% vs 54%, t(40) = 4.60; Bonferroni, p < .005). Also, it was found that they accepted the conjunction ‘not-A & B’ more frequently than the conjunctions ‘A & B’ and ‘not-A & not-B’ (93% vs 3%, t(40) = 15.43; 93% vs 5%, t(40) = 16.97; Bonferroni, p < .005).


<table>
<thead>
<tr>
<th></th>
<th>B, if not-A</th>
<th>If not-A, then B</th>
<th>B, except if A</th>
<th>Except if A, B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A and B</td>
<td>5</td>
<td>3</td>
<td>20</td>
<td>54</td>
</tr>
<tr>
<td>A and not-B</td>
<td>46</td>
<td>54</td>
<td>63</td>
<td>32</td>
</tr>
<tr>
<td>Not-A and B</td>
<td>93</td>
<td>93</td>
<td>76</td>
<td>41</td>
</tr>
<tr>
<td>Not-A and not-B</td>
<td>15</td>
<td>5</td>
<td>7</td>
<td>26</td>
</tr>
</tbody>
</table>

They accepted the conjunction ‘A & not-B’ as frequently as the conjunction ‘not-A & B’ in the exceptive conditional ‘B except if A’ (63% vs 76%, t(40) = 1.40, p = .17; Bonferroni, p < .005). Also, they accepted the conjunction ‘not-A & B’ more frequently than the conjunctions ‘A & B’ and ‘not-A & not-B’ (76% vs 20%, t(40) = 4.83; 76% vs 7%, t(40) = 7.07; Bonferroni, p < .005) and they accepted the conjunction ‘A & not-B’ more frequently than the conjunctions ‘A & B’ and ‘not-A & not-B’ (63% vs 20%, t(40) = 3.78; 63% vs 7%, t(40) = 6.05; ES = .69; Bonferroni, p < .005). Finally, there were no reliable differences between conjunctions when the conditional was ‘except if A, B’ (Bonferroni, p > .05, in all cases). There was no main effect of type of conditional, F (3, 38) = .626, MSE =
0.060, \( p = .63 \), \( \eta^2 = .014 \), but there was a main effect of type of conjunction, \( F(2.015, 50.593) = 43.43, \) \( MSE = .39, p < .001, \eta^2 = .52 \).

Experiment 2 suggests that participants reason with the conditional ‘B, if not-A’ and ‘if not-A, then B’ by initially envisaging only one initial possibility that corresponds to the fact ‘not-A and B’, as suggested by model theory (Johnson-Laird & Byrne, 2002). Also, Experiment 2 shows that participants reason with the exceptive conditional ‘B except if A’ by initially envisaging two complete possibilities: one that corresponds to the probable fact ‘not-A & B’ and a second that corresponds to the exceptive possibility ‘A & not-B’. In brief, the fact that most participants accepted the conjunctions ‘A & not-B’ and ‘not-A & B’ more frequently than other conjunctions (‘A & B’ and ‘not-A & not-B’) in ‘B except if A’, corroborates our predictions that both possibilities can be obtained from the initial representation. This result runs counter to the predictions that we derived from Suppositional theory, which proposes that people think about one possibility. Unexpectedly, our predictions about the exceptive conditional ‘except if A, B’ were not confirmed. We predicted that participants should have accepted the possibilities ‘A & not-B’ and ‘not-A & B’ more than the possibilities ‘A & B’ and ‘not-A & not-B’. However, neither of these predictions was fulfilled. As in Experiment 1, we suggest that participants did not understand the meaning of this connective, because in the Spanish language the clauses with ‘except if’ most commonly follow the main clauses (‘B except if A’) and conditionals with pre-posed clauses (‘except if A, B’) are very rare and unusual.

**GENERAL DISCUSSION**

In this paper we provide for the first time (to our knowledge) new evidence on the psychological interpretation of the conditional *except if* in two reasoning tasks. The results of Experiments 1 and 2 corroborate the idea that the linguistic expression *except if* exerts certain forms of modulation on the meaning of the conditional and determines the inferences that individuals draw in an inference task (Experiment 1) and truth table task (Experiment 2).

Experiment 1 showed that there were reliable differences between MP and MT inferences in the conditionals ‘A, if not- B’ and ‘if not-A, then B’ but not in the exceptive conditionals ‘B except if A’ and ‘except if A, B’.

The absence of any difference between MP and MT in the conditional ‘B except if A’ rules out alternative theories of reasoning, such as formal rule-based theories (Braine & O’Brien, 1998; Rips, 2004) and Suppositional
Reasoning with Exceptive Conditionals

theory (Evans, 2007; Evans & Over, 2004). From the point of view of these theories, MT inferences should be more difficult than MP inferences. However, our results can be explained by model theory (Johnson-Laird & Byrne, 1991, 2002), which claims that participants should make more MP than MT inferences in the conditionals if *not*, *then*, because the MP inference can be made from the initial models, but to obtain the MT inferences they have to flesh out the models, that is, make the implicit model explicit. However, model theory could explain the absence of reliable differences between MP and MT inferences in ‘B except if A’, due to the fact that participants have two initial mental models (‘A & not-B’ and ‘not-A & B’) for this connectives. As a result of this dual initial representation, they can make both the MP and the MT inferences. The absence of reliable differences between AC and DA inferences lends support again to the idea that people have two initial mental models with the conditional except if. On the other hand, Experiment 1 showed that participants made fewer MT inferences than MP inferences with the conditional ‘if not-A, then B’ (or ‘B, if not-A’), and they made more AC than DA inferences with the conditional ‘B, if not-A’; however, the differences were not significant with the conditional ‘if not-A, B’, although the differences were in the predicted direction. We suggest that the absence of difference between AC and DA inferences for the conditional ‘B, if not-A’ is due to the fact that the advantage of being part of the initial model of AC inference can be counteracted by backward processing.

Experiment 1 also showed that participants presented a high percentage of asymmetric responses with the conditional ‘except if A, B’. These data could suggest that participants had problems understanding the meaning of the expression when the clause except if was pre-posed. Different studies have shown that the postposed clause in complex conditional connectives (such as except if) occupy the second position (such as ‘B except if A’) in the conditional construction in English (Traugott, 1997; Dancygier, 1998) and in Spanish (Montolío, 2000). We suggest that when people have to reason with unusual expressions such as ‘except if A, B’, they may use a shortcut strategy that consists of matching the two terms that appear in the statement.

Experiment 2 showed that in the conditionals ‘B, if not-A’ and ‘if not-A, then B’ participants tended to accept the possibility ‘not-A & B’ more frequently than other possibilities (‘A & B’, ‘A & not-B’ and ‘not-A & not-B’). These results confirm the idea that people accept the possibility that corresponds to an initial mental model (‘not-A & B’) more frequently than the one that corresponds to an implicit model (‘A & B’ and ‘A & not-B’) or false model (not-A & not-B). Also, Experiment 2 showed in the conditional
‘B except if A’, that participants accepted the possibilities ‘A & not-B’ and ‘not-A & B’ more frequently than the possibilities ‘A & B’ and ‘not-A & not-B’. This finding supports the idea that when participants reason with the conditional ‘B except if A’ they envisage two possibilities, i.e. one that corresponds to the probable fact ‘not-A & B’ and a second that corresponds to the exceptive possibility ‘A & not-B’.

The results from Experiment 2 are inconsistent with the minimalist hypothesis (Sloutsky & Goldvarg, 2004), which claims that people tend to construct only single models for connectives. Also, our results run counter to the predictions that we derived from Suppositional theory, which proposes that people think about possibilities that contain the true antecedent of the assertion but not the false antecedent (Evans et al., 2005). In keeping with this idea, it is predicted that there should be differences between the possibilities ‘A & not-B’ and ‘not-A & B’ in the conditional ‘B except if A’. However, no differences were found between these two possibilities. This result may also be difficult to explain using theories of formal rules of inference (Braine & O’Brien, 1998; Rips, 1994) and theories based on domain-specific rules of inference (Fiddick, Cosmides, & Tooby, 2000; Gigerenzer & Hug, 1992; Holyoak & Cheng, 1995). These theories do not provide the means by which predictions can be derived about whether certain conjunctions should be accepted more frequently by conditionals, and other conjunctions should not. However, our predictions about ‘except if A, B’ were not confirmed. We predicted that participants would accept the possibilities ‘A & not-B’ and ‘not-A & B’ more than the possibilities ‘A & B’ and ‘not-A & not-B’. However, neither of these predictions was fulfilled. As in Experiment 1, this result could suggest that participants did not understand the meaning of this connective, because in the Spanish language the clauses with except if most commonly follow the main clauses (‘B except if A’) and conditionals with pre-posed clauses (‘except if A, B’) are very rare and unusual.

Results from Experiments 1 and 2 are consistent with other results obtained by priming methodology (Espino et al., 2013; Gómez-Veiga et al., 2012). Espino et al. (2013) found that the possibility ‘A & not-B’ was primed by ‘B except if A’ when compared to ‘B, if not-A’, whereas there were no differences between the priming effects of the two conditionals on the reading time of ‘not-A & B’. Gómez-Veiga et al. (2012) found that the possibility ‘not-A & B’ was primed by ‘unless A, B’ when compared to ‘if A, not-B’, whereas there were no differences between the priming effects of the two conditionals on the reading time of ‘A & not-B’. Also, there were no differences between different conditional forms when participants read the test sentences ‘A & B’ and ‘not-A & not-B’. These results could suggest
that the exceptive conditionals except if and unless could be semantically analogous. From our point of view, unless and except if are semantically equivalent in the sense that both have bi-conditional meaning (Gómez-Veiga et al., 2012; Espino et al., 2013; Montolío, 1999, 2000), but the mental representation underlying the two conditionals could be different, because the conditional unless requires the subjunctive mood in the subordinate clause, while the conditional except if requires the indicative. The indicative mood is used to express factual information, certainty and objectivity, while the subjunctive mood conveys wishes, conjectures and uncertainty. We suggest that people could use except if to express factual information, certainty, and objectivity in an exceptive context while they could use unless to express conjectures and uncertainties in an exceptive context. A fruitful avenue for future research would be to examine the mental representation underlying both negative exceptive conditionals.

One important aspect of our research was that we found convergent results with two different reasoning tasks. Some authors (e.g., Barrouillet, Gauffroy, & Lecas, 2008) have claimed that these two tasks show different aspects of the representation in reasoning. The first is used more frequently by the mental model theorist and the second by the suppositional theorist. According to Evans (2007), the results obtained by the truth table task is at odds with mental model theory, but provides the main evidence for Suppositional theory. However, the result of both tasks in the present study are congruent with mental model theory and difficult to reconcile with Suppositional theory (Evans, 2007). Recently, Vergauwe, Morsanyi, Dagry, & Barrouillet (2013), using a truth table task, found that the entire pattern of response times almost perfectly fitted the predictions issuing from mental model theory. They concluded with the assertion that contrary to Evans’ (2007) claim, ‘the defective truth table task is not a piece of evidence in favor of Suppositional theory’ (Vergauwe et al., 2013, page 181).

To summarise, the view of the model theory that a key difference in understanding and reasoning with different conditionals lies in the possibilities that people keep in mind is confirmed. Our results confirm the hypothesis that people initially interpret the conditional ‘B except if A’ in terms of a dual representation due to the modulation effect. Also, Experiments 1 and 2 support the claim that ‘B except if A’ is different than ‘B, if not-A’ and also different than ‘if not-A, then B’ (Dancygier & Sweetser, 2005; Declerck & Reed, 2000). Finally, these results refute the hypothesis that ‘B except if A’ is semantically equivalent to ‘except if A, B’ (Declerck & Reed, 2000).
RESUMEN

Razonamiento con condicionales exceptivos: El caso de “excepto si”. En esta investigación presentamos una adaptación de la teoría de modelos mentales (Johnson-Laird & Byrne, 1991, 2002) que trata de explicar cómo las personas razonan con el condicional excepto si. Se realizaron dos experimentos que mostraron que el condicional exceptivo excepto si ejerce ciertos tipos de modulación semántica y determina las inferencias que los individuos elaboran en una tarea de inferencia (Experimento 1) y de tablas de verdad (Experimento 2). En el Experimento 1 encontramos que no hay diferencias entre el porcentaje de Modus Ponens y Modus Tollens en el condicional excepto si pero sí son significativas para el condicional si no, entonces. En el Experimento 2, los participantes seleccionaron más frecuentemente las posibilidades ‘A y no-B’ y ‘no-A y B’ que las posibilidades ‘A y B’ y ‘no-A y no-B’ en el condicional ‘B excepto si A’, mientras que con el condicional ‘B si no A’ y ‘si no A, B’ seleccionaron más frecuentemente la posibilidad ‘no-A y B’ que las otras posibilidades (‘A y no-B’, ‘A y B’, ‘no-A y no-B’). Las implicaciones de estos resultados se discuten en el contexto de las teorías psicológicas y lingüísticas recientes referentes al significado de excepto si.

REFERENCES


(Manuscript received: 27 November 2013; accepted: 19 November 2014)