

## **UC Merced**

### **Proceedings of the Annual Meeting of the Cognitive Science Society**

#### **Title**

The elusive oddness of or-introduction

#### **Permalink**

<https://escholarship.org/uc/item/1079f1nh>

#### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 39(0)

#### **Authors**

Cruz, Nicole

Over, David

Oaksford, Mike

#### **Publication Date**

2017

Peer reviewed

# The elusive oddness of or-introduction

Nicole Cruz<sup>1,2</sup> (ncruzd01@mail.bbk.ac.uk), David Over<sup>3</sup> (david.over@durham.ac.uk), & Mike Oaksford<sup>2</sup> (m.oaksford@bbk.ac.uk)

<sup>1</sup>Laboratoire CHArt (EPHE), 4-14, rue Ferrus, 75014 Paris, France

<sup>2</sup>Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK

<sup>3</sup>Department of Psychology, Durham University, South Rd, Durham DH1 3LE, UK

## Abstract

The inference of or-introduction,  $p$ , therefore  $p$  or  $q$ , is fundamental in classical logic and probability theory. Yet traditional research in the psychology of reasoning found that people did not endorse this inference as highly as other one-premise valid inferences. A radical response to this finding is to claim that or-introduction is in fact invalid. This response is found in the recent revision of mental model theory (MMT). We argue that this revision of the theory leads to a number of logical problems and counterintuitive consequences for valid inferences, and present an experiment extending recent studies showing that people readily accept or-introduction under probabilistic instructions. We argue for a pragmatic explanation of why the inference is sometimes considered odd. The inference is not odd when people reason from their degrees of belief.

**Keywords:** or-introduction; reasoning; mental models; probabilistic approach

## The New Paradigm and earlier MMT

There has been a paradigm shift in the psychology of reasoning (Oaksford & Chater, 2013; Over, 2009), from binary approaches focussed on drawing conclusions from arbitrary assumptions, to Bayesian and probabilistic accounts focussed on people's degrees of belief and belief revision and updating in reasoning.

## The probabilistic approach

A central position in the probabilistic approach is that most reasoning in both everyday life and science takes place under uncertainty. This uncertainty cannot be captured in classical binary logic, but it can be in probability theory (Adams, 1998; Coletti & Scozzafava, 2002).

A basic hypothesis in this new approach is that people's degree of belief in a conditional statement,  $P(\text{if } p \text{ then } q)$ , does not correspond to the probability of the material conditional of classical logic (which is equivalent to  $\text{not-}p$  or  $q$ ), but instead to the conditional probability,  $P(q|p)$ . The proposal is that people arrive at this conditional probability by performing a *Ramsey test*, a mental simulation in which they hypothetically add  $p$  to their beliefs, make any changes necessary to preserve consistency, and judge the probability of  $q$  on this basis (Evans & Over, 2004; Ramsey, 1929/1990; Stalnaker, 1968). The identity  $P(\text{if } p \text{ then } q) = P(q|p)$  is generally called *The Equation* (Edgington, 1995) and has received strong empirical support (Evans, Handley,

Neilens, & Over, 2007; Evans, Handley, & Over, 2003; Oberauer, Geiger, Fischer, & Weidenfeld, 2007; Oberauer & Wilhelm, 2003; Fugard, Pfeifer, Mayerhofer, & Kleiter, 2011; Barrouillet & Geauffroy, 2015).

## Probabilities in earlier MMT

Earlier versions of mental model theory (MMT) proposed that people reason by creating a mental representation of the logical possibilities in which the premises of an inference are true (e. g.  $p$  &  $q$  is true in one possibility: that in which both  $p$  and  $q$  are true; whereas  $p$  or  $q$  is true in three possibilities: when  $p$  is true and  $q$  false, when  $p$  is false and  $q$  true, and when  $p$  and  $q$  are both true). Each of these possibilities is called a model. People then eliminate any models of the premises that contradict one another (e. g. if Premise 1 of an inference is  $p$  or  $q$  and Premise 2 is  $\text{not-}p$ , then people eliminate the two models of Premise 1 in which  $p$  is true). Finally, people formulate an informative conclusion based on any models remaining after eliminating inconsistencies. It was further held that people make errors in reasoning because they tend to represent only what is true in a model, and to leave implicit what is false, and because they tend to leave implicit and then forget entire models.

MMT was originally formulated within the binary approach to reasoning. Hence it focussed on the truth or falsity of a statement, given the truth of some other statements, and proposed the core meaning of conditionals to correspond to the material conditional (Johnson-Laird & Byrne, 1991, 2002).

However, MMT was early on extended to reasoning with extensional probabilities, representing these as proportions of models or numerical "tags" on models (Giroto & Johnson-Laird, 2004, 2010; Johnson-Laird, Legrenzi, Giroto, Legrenzi, & Caverni, 1999; c. f. Geiger & Oberauer, 2010) in a way consistent with the rules of probability theory.

MMT has also been recently extended to subjective probabilities (Khemlani, Lotstein, & Johnson-Laird, 2014). However, a problem with this account is the lack of clarity in its computational level specification. For example, it proposes that people intuitively grasp that the "logical relation"  $p$  or  $\text{not-}p$  has a probability of 1, but also that people intuitively compute  $P(p \text{ or } q)$  by taking the average of  $p$  and  $q$  (cf. Juslin, Nilsson, & Winman, 2009) – even though the logical connective is the same in both cases. The account also provides no means for computing correct

conditional probabilities, and can therefore not account for their occurrence. This contrasts with the earlier proposal in MMT of the *subset principle* for computing extensional conditional probabilities (Johnson-Laird et al., 1999), which could account for both errors and normative responses.

### New MMT

Johnson-Laird and colleagues recently proposed a radical revision of MMT that claims to integrate logic further with probability (Johnson-Laird, Khemlani, & Goodwin, 2015). The revision changes the meanings of conditionals, conjunctions and disjunctions. These statement types are still represented using the same models as before, but whereas in previous versions of MMT a statement was true when at least one of its models was true, in the new version of the theory a statement is true when *all* of its models are *possible* (Johnson-Laird et al., 2015).

A positive consequence of this revision, in our view, is that the *paradoxes of the material conditional* are now considered invalid. For example, the inference *q, therefore if p then q* is considered invalid because the model for *q* does not establish that all three models of the material conditional are possible. But the revision creates a number of logical problems and counterintuitive consequences for other inferences (Baratgin, Douven, Evans, Oaksford, Over, & Politzer, 2015; Over & Cruz, in press).

### Logical problems with new MMT

#### "Possible", "true", and "valid"

Johnson-Laird et al. (2015) argue that a statement is true when all of its models are possible. But it is not clear what is meant by "possible". It cannot be logical possibility, because logically the four combinations of the truth and falsity of *p* and of *q* (*p & q*, *p & not-q*, *not-p & q*, *not-p & not-q*) are always possible unless they contain a contradiction (Baratgin et al., 2015). Moreover, with logical possibility the new version of MMT would imply that the tautology *p or not-p* is false because the *p & not-p* model is not possible (Baratgin et al., 2015).

Yet a narrower notion of possibility does not seem to work either. This can be seen if we apply the idea that a statement is true whenever all of its models are possible to statements that have a single model. The theory then implies that a statement *p* is true when it is possible. But it can be possible, and readily conceivable, for us to sleep a little longer tomorrow, and yet be false if we wake up early instead. Truth does not follow from mere possibility, no matter how it is defined (Over & Cruz, in press).

A further problem arises when the notion of truth in new MMT is used to assess the validity of an inference. Johnson-Laird et al. (2015) continue to define an inference as logically valid when its conclusion is true in every case in which its premises are true. If a statement is true when all of its models are possible, then by implication an inference is valid when the truth of the premises establishes that all models of the conclusion are possible. This formulation of

the concept of validity is of course difficult to understand without a clear definition of "possible". But one way of operationalising it could be as follows. An inference is valid in new MMT when the models of the premises contain all the models of the conclusion. This operationalisation renders the paradoxes invalid. But it leads to counterintuitive conclusions for other inferences. For example, it implies that the inference *p or q, therefore p* is valid, even though it is counterintuitive and invalid in classical and probabilistic logics. At the same time, the account implies that the inference *not-p, therefore not-(p & q)* is invalid, but this is an intuitive inference to make, and it is valid in classical and probabilistic logics.

#### Or-introduction

In what follows we focus on the inference of or-introduction, *p, therefore p or q*. This inference is valid in classical logic and in the probabilistic approach because it is incoherent to judge that  $P(p) > P(p \text{ or } q)$ . It was also valid in previous versions of MMT.

Past studies using binary instructions (asking participants to assume the premises to be true, and then to judge whether the conclusion also had to be true) found that people accept the inference less frequently than other valid inferences (Braine, Reiser, & Romain, 1984; Orenes & Johnson-Laird, 2012; Rips, 1983). But the probabilistic approach and previous versions of MMT agreed that the lower acceptance rate can be explained through pragmatic factors: *p* is a stronger, more informative statement than *p or q*, and so it is pragmatically infelicitous to assert *p or q* when one has enough information to assert *p* (Grice, 1989). Orenes & Johnson-Laird (2012) specified this position further, suggesting that the pragmatic infelicity of the inference comes from the fact that the conclusion *p or q* includes a model in which the premise *p* is false. We agree with Grice (1989) that it is potentially misleading to assert *p or q* in a conversation, suggesting that *p* is possibly false, after inferring *p or q* from *p* (Gilio & Over, 2012).

However, in the new version of MMT or-introduction is considered invalid for the same reason as the paradoxes: the model for *p* does not establish that the three models for the disjunction are possible. This revision does not take into account the more recent finding that under probabilistic instructions (asking participants for their degree of belief in the premise, and for their degree of belief in the conclusion given their degree of belief in the premise) or-introduction is accepted to a high degree, indistinguishable from that of other, uncontroversially valid inferences (Cruz, Baratgin, Oaksford, & Over, 2015; Politzer & Baratgin, 2016). According to the probabilistic approach, people accept the inference under probabilistic instructions because pragmatic constraints related to what is asserted in a conversation tend to be eliminated or reduced when people are asked directly for their subjective beliefs. Conversational principles about not misleading our hearers (Grice, 1989) do not apply when we are making inferences from our own beliefs as premises to further beliefs in a subjective mental process.

The assumed invalidity of or-introduction also disables central logical and probabilistic principles. For example, if or-introduction is invalid, then so is a version of the inference in which the disjunction is "packed": *p, therefore superset of p*. For example, *there is tea, therefore there is tea or coffee* can be paraphrased as *there is tea, therefore there is a hot beverage*. Or-introduction thus enables us to establish basic set-subset relations. If people are unable to establish such relations, then the MMT account of reasoning with categorical syllogisms breaks down. Or-introduction is also used in the proofs of many fundamental theorems of probability theory, such as the theorem of total probability,  $P(p) = P(p \ \& \ q) + P(p \ \& \ \text{not-}q)$ , which is itself derived from the fundamental logical principle that *p* is equivalent to *(p & q) or (p & not-q)*. MMT cannot integrate logic and probability theory while implying that such principles are invalid.

In what follows we analyse in more detail the relation between or-introduction and two further inferences: and-elimination, *p & q, therefore p*, and or-MP: *if p or q then r, p, therefore r*.

### And-elimination = or-introduction

And-elimination appears to be valid in the new version of MMT, because the model of the premise contains all the models of the conclusion. But the validity of and-elimination implies the validity of or-introduction, and vice versa, as follows:

*p & q, therefore p*  
*not-p, therefore not-(p & q)* (by reductio ad absurdum)  
*not-p, therefore not-p or not-q* (by de Morgan)  
*p, therefore p or q* (by substitution of terms)

Mental model theory could argue that it does not accept this proof because one or more of the rules used in the derivation are itself invalid in the theory. But the invalidity of such elementary logical rules would have counterintuitive consequences for a wide range of further inferences.

### Or-MP: or-introduction through the back door

The inference *if p or q then r, p, therefore r* can be called or-MP because it is the short form of a two-step inference, in which one first uses or-introduction to infer *p or q* from *p*, then then uses *p or q* together with *if p or q then r* to infer *r* through the inference of modus ponens (MP).

Under binary instructions or-MP is endorsed to a degree at least as high as MP (Rips, 1983). The inference also appears to be valid in new MMT because the models of the premises contain all the models of the conclusion. But as outlined above, or-MP includes or-introduction as a component.

Followers of MMT might reply that, in the new version of the theory, or-MP is valid directly without the intermediate step of or-introduction. But the validity of or-MP also implies the validity of or-introduction in a direct way. If we substitute *p or q* for *r*, the resulting inference is *if p or q*

*then p or q, p, therefore p or q*. The first premise is a tautology, which always holds and does not have to be assumed, and the rest is explicit or-introduction (c. f. Over & Cruz, in press).

We conducted an experiment to test people's intuitions about the validity of or-introduction, and its relation to and-elimination and or-MP, using probabilistic instructions.

## Method

### Participants

A total of 121 participants from English speaking countries completed the experiment through the online platform Prolific Academic. After removing cases that failed a test question or included trial reaction times of 3 sec or less, 112 participants remained for analysis. They had a mean age of 29 years (range: 18-73), and a varied formal-educational background. All indicated having at least good English language skills. Participants' percentage rating of task difficulty was on average 48%.

### Design and materials

Participants were shown the 6 inferences of Table 1 (two further inferences investigating other questions are not discussed here due to space constraints). Inferences 1 to 5 were presented three times, with three different premise probabilities (100%, 80%, and 60%). Inference 6 was only presented with a premise probability of 100% because one of its premises is a tautology. Premise probability was varied with the aim of generalising the results to different premise probabilities, and was not associated with particular predictions. For each trial, participants' task was to judge how likely the conclusion of the inference can be, given the likelihood of the premise.

Table 1. The inferences investigated.

Name	Form
1 or-introduction	$p \therefore p \text{ or } q$
2 and-elimination	$p \ \& \ q \therefore p$
3 Paradox 1	$q \therefore \text{if } p \text{ then } q$
4 Paradox 2	$\text{not-}p \therefore \text{if } p \text{ then } q$
5 or-MP (a)	$\text{if } p \text{ or } q \text{ then } r, p \therefore r$
6 or-MP (b)	$\text{if } p \text{ or } q \text{ then } p \text{ or } q, p \therefore p \text{ or } q$

Note. " $\therefore$ " = "therefore".

The experiment involved three comparisons. The first was between inference 1 (or-introduction) and inference 2 (and-elimination). The probabilistic approach predicts people will give similar ratings to these two inferences because the validity of one implies the validity of the other, and asking directly for people's degree of belief in the conclusion is expected to reduce pragmatic factors that may have led to lower acceptance rates of or-introduction using binary instructions (Orenes & Johnson-Laird, 2012; Rips, 1983). In contrast, new MMT would predict inference 2 to be

accepted to a high degree but inference 1 to be rejected, because inference 2 is valid but inference 1 is invalid in the new version of the theory.

The second comparison was between inference 1 (or-introduction) and inferences 3 and 4 (the paradoxes of the material conditional). The probabilistic approach predicts people will accept inference 1 to a higher degree than inferences 3 and 4 because the first is valid but the latter two invalid. In contrast, new MMT would predict that people will reject all three inferences to a similar degree, because they are considered invalid for the same reason in the theory.

The third comparison was between inference 1 (or-introduction) and inferences 5 and 6 (or-MP). The probabilistic approach predicts people will endorse inferences 1, 5 and 6 to a similar degree, because they are all valid. New MMT would predict inference 5 to be accepted as valid (by assuming it to be computed in a direct way without the intermediate step of or-introduction), but inferences 1 and 6 to be rejected because or-introduction is invalid on their account.

On each trial participants saw an inference embedded in a pseudo-naturalistic context story. The context stories changed on every trial, and were randomly allocated to the inferences for each participant. The order of occurrence of the inferences was also varied randomly for each participant. With 8 inferences (two not reported here) and 3 probabilities (and inference 6 only being paired with a probability of 100%), there were 22 trials overall, plus two control trials to check whether participants were paying attention.

## Procedure

After going through the instructions and three practice trials involving different inferences to those in Table 1, participants worked through the 24 trials of the experiment. They then provided demographical information and indicated whether they had taken part seriously. The final page provided debriefing information. The experiment took on average 13.2 minutes to complete.

## Results and discussion

To compare the above predictions of the probabilistic approach and the revised version of MMT, two linear mixed model analyses were performed. The procedure for model construction followed the recommendation of Barr, Levy, Scheepers, & Tily (2013) of implementing the maximum possible random effects structure justified by the design. The models included a random intercept for participants, but random effects for material could not be included because the random allocation of materials to inferences had as a consequence that there were not enough repetitions of the same type of material within each cell of the design. Predictor variables were centred around their grand mean to avoid problems of multicollinearity when including interaction terms. Comparisons of the main *F*-test results with likelihood-ratio tests led to the same pattern of

significant and non-significant effects. Effect sizes were calculated using the formulas suggested by Snijders and Bosker (2012), requiring the use of ML as opposed to reML as estimation method. The results are displayed in Figure 1.

### Analysis 1: Inferences 1 to 5

We first fitted an overall model with inference (1 to 5) and probability (100%, 80%, 60%) as independent variables<sup>1</sup>. Judgments of conclusion probability increased with increasing premise probability,  $F(2, 1568) = 118.76, p < .001, \eta_p^2 = .074$ . Mean probability judgments differed between inferences,  $F(4, 1568) = 269.09, p < .001, \eta_p^2 = .334$ . The size of the effect of premise probability also differed between inferences,  $F(8, 1568) = 13.06, p < .001, \eta_p^2 = .055$ .

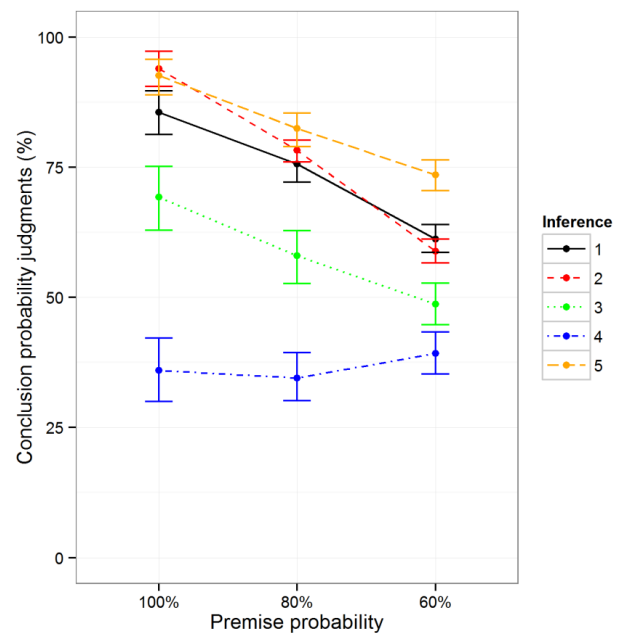


Figure 1. Judgments of conclusion probability for inferences 1 to 5 as a function of premise probability. Error bars show 95% CIs.

<sup>1</sup>Following the notation of Snijders & Bosker (2012), the equation for measurement *i* of participant *j* was given by:

$$Y_{ij} = \gamma_{00} + \gamma_{10}x_{1ij} + \gamma_{20}x_{2ij} + \gamma_{30}x_{1ij}x_{2ij} + u_{0j} + \epsilon_{ij}$$

This random coefficients model had 17 parameters: 1 for the fixed effect of the intercept, 4 for the fixed effect of inference, 2 for the fixed effect of premise probability, 8 for the fixed interaction between inference and premise probability, 1 for the variance of the intercept, and 1 for the residual variance. The fact that the predictors were centred is not represented in the equation due to space constraints. The equations for the other linear mixed models computed in the analyses followed the same principles, but are not reported due to limitations of space.

Figure 1 suggests that the interaction between inference and premise probability can be traced back to the lack of an effect of premise probability for inference 4. Follow-up analyses showed that there was indeed no effect of premise probability for inference 4,  $F(2, 224) = 1.06, p = .35, \eta_p^2 = .005$ . However, the size of the effect of premise probability still varied between inferences 1, 2, 3 and 5,  $F(2, 1232) = 189.60, p < .001, \eta_p^2 = .017$ .

Follow-up analyses to the effect of inference showed that there was no significant difference in probability judgments for inference 1 ( $M = 75.14, SE = 1.34$ ) and inference 2 ( $M = 78.02, SE = .82$ ) ( $F(1, 560) = 3.23, p = .073, \eta_p^2 = .005$ ). This is in accordance with the predictions of the probabilistic approach, because or-introduction and and-elimination can be derived from one another as valid inferences. The result is at odds with new MMT, which predicts the second to be valid but the first invalid.

Judgments for inference 1 were higher than those for inference 3 ( $M = 59.66, SE = 1.77$ ) ( $F(1, 560) = 70.028, p < .001, \eta_p^2 = .085$ ) and than those for inference 4 ( $M = 37.58, SE = 1.70$ ) ( $F(1, 560) = 424.164, p < .001, \eta_p^2 = .371$ ). This is again in accordance with the probabilistic approach, for which or-introduction is valid, but the paradoxes of the material conditional are not. It goes counter to new MMT, in which the three inferences are invalid for the same reason.

Judgments for inference 5 ( $M = 83.85, SE = 1.03$ ) were slightly higher than those for inference 1 ( $F(1, 560) = 33.47, p < .001, \eta_p^2 = .044$ ) and than those for inference 2 ( $F(1, 672) = 14.990, p < .001, \eta_p^2 = .022$ ). Taken by itself this finding is in accordance with both the probabilistic approach and new MMT, because both predict the inference to be valid. Small differences in the acceptance of valid inferences are not a problem for either theory, as long as the difference between responses to valid and responses to invalid inferences is larger, as Figure 1 clearly corroborates. The slightly higher acceptance of or-MP than of or-introduction and and-elimination is in accordance with the fact that or-MP includes MP as a component, and MP tends to be endorsed at ceiling. Responses to inference 5 become more consequential to the questions investigated here when compared to those of inference 6.

## Analysis 2: Inferences 1, 5, and 6

We next fitted a model with inference (1, 5, 6) as the independent variable, using responses for a premise probability of 100%. Judgments for inference 1 ( $M = 85.60, SE = 1.98$ ) were again slightly lower than those for inference 5 ( $M = 92.63, SE = 1.98$ ) ( $F(1, 112) = 8.34, p = .005, \eta_p^2 = .027$ ) and than those for inference 6 ( $M = 91.32, SE = 2.03$ ) ( $F(1, 112) = 4.69, p = .032, \eta_p^2 = .019$ ). Judgments for inference 5 did not differ from those for inference 6 ( $F(1, 112) = .31, p = .58$ ). Thus version (a) of or-MP (*if p or q then r, therefore r*) and version (b) of the inference (*if p or q then p or q, p, therefore p or q*) were endorsed to the same degree, even though version (b) explicitly contains or-introduction as a component. This is

in accordance with the probabilistic approach, under which the two inferences are equivalent, but at odds with new MMT, which would predict version (b) of the inference to be rejected because or-introduction is invalid in its account.

An interesting, not anticipated finding concerns the pattern of results for the paradoxes (inferences 3 and 4). For inference 4 there was no effect of premise probability, and judgments of conclusion probability were consistently low. Judgments for inference 3 were also clearly lower than those for inferences 1, 2, and 5, but they did covary positively with premise probability.

A first account of this difference could be as follows. In the case of inference 4, *not-p, therefore if p then q*, premise and conclusion contain no elements in common, and so the fact that the premise is uninformative about the conclusion is clearly apparent. Without any information about the conclusion, people assign a low probability to it, expressing that it does not follow from the premise. In the case of inference 3, *q, therefore if p then q*, the premise is also uninformative about the conclusion, and so any response is coherent. But in the absence of further information, it is reasonable to infer that a given probability of *q* will remain invariant under the assumption of *p*. It therefore makes sense for responses to covary positively with premise probability.

## General discussion

We investigated the inference of or-introduction and its relation to and-elimination and or-MP. Earlier research using binary instructions had found or-introduction to be accepted less frequently than and-elimination (Rips, 1983). The recent revision of MMT argues that or-introduction is in fact invalid. But the assumptions of this revision have inconsistencies and counterintuitive consequences for other inferences. We extended recent findings (Cruz et al., 2015; Polizer & Baratgin, 2016) using probabilistic instructions and found that or-introduction is accepted to a high degree, indistinguishable from that of and-elimination. People's responses to or-MP were slightly higher than those for or-introduction and and-elimination, even though or-MP contains or-introduction as a component. These findings are in accordance with the predictions of a Bayesian approach to the study of reasoning, but not with those of new MMT.

With the exception of the paradoxes, the findings could also have been accounted for in earlier versions of MMT concerned with extensional probabilities (Giroto & Johnson-Laird, 2004, 2010; Johnson-Laird et al., 1999; c. f. Orenes & Johnson-Laird, 2012). These earlier formulations converged with the probabilistic approach in holding that or-introduction is valid, but is sometimes odd for pragmatic reasons. The results of this experiment provide further evidence for the pragmatic explanation. The findings also suggest that people tend to reason using or-introduction in a more logical and less biased way under probabilistic than under binary instructions. People's inferences from their own degrees of belief to further degrees of belief do not seem to be governed by the conventions of conversation for

speakers and hearers in open discussions, but rather by the Bayesian principles of belief revision and updating, as proposed in the new paradigm in the psychology of reasoning.

### Acknowledgments

We are grateful to the Mexican National Council of Science and Technology (CONACYT) for supporting this research.

### References

- Adams, E. (1998). *A primer of probability logic*. Stanford, US: CLSI publications.
- Baratgin, J., Douven, I., Evans, J. St. B. T., Oaksford, M., Over, D. E., & Politzer, G. (2015). The new paradigm and mental models. *Trends in Cognitive Sciences*, 19(10), 547-548.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and language*, 68, 255-278.
- Barrouillet, P., & Gauffroy, C. (2015). Probability in reasoning: A developmental test on conditionals. *Cognition*, 137, 22-39.
- Braine, M. D. S., Reiser, B. J., & Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. *Psychology of Learning and Motivation*, 18, 313–337.
- Coletti, G., & Scozzafava, R. (2002). Probabilistic logic in a coherent setting. Dordrecht, NL: Kluwer.
- Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology*, 6, 192.
- Edgington, D. (1995). On conditionals. *Mind*, 104, 235-329.
- Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 321-335.
- Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford, UK: Oxford University Press.
- Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 635-648.
- Geiger, S. M., & Oberauer, K. (2010). Towards a reconciliation of mental model theory and probabilistic theories of conditionals. In M. Oaksford, & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (289-308). Oxford, UK: Oxford University Press.
- Gilio, A., & Over, D.E. (2012). The psychology of inferring conditionals from disjunctions: A probabilistic study. *Journal of Mathematical Psychology*, 56, 118-131.
- Giroto, V., & Johnson-Laird, P. N. (2004). The probability of conditionals. *Psychologia*, 47, 207-225.
- Giroto, V., & Johnson-Laird, P. N. (2010). Conditionals and probability. In M. Oaksford, & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (103-116). Oxford, UK: Oxford University Press.
- Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, US: Harvard University Press.
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hillsdale, US: Erlbaum.
- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109(4), 646-678.
- Johnson-Laird, P. N., Khemlani, S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, 19(4), 201-214.
- Johnson-Laird, P. N., Legrenzi, P., Giroto, V., Legrenzi, M. S., & Caverni, J. P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, 106, 62-88.
- Juslin, P., Nilsson, H., & Winman, A. (2009). Probability theory, not the very guide of life. *Psychological Review*, 116, 856–874.
- Khemlani, S. S., Lotstein, M., & Johnson-Laird, P. N. (2015). Naive probability: Model-based estimates of unique events. *Cognitive Science*, 39, 1216-1258.
- Oaksford M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking & Reasoning*, 19, 346-379.
- Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 680-693.
- Orenes, I., & Johnson-Laird, P. N. (2012). Logic, models, and paradoxical inferences. *Mind & language*, 27(4), 357-377.
- Over, D. E. (2009). New paradigm psychology of reasoning. *Thinking & Reasoning*, 15, 431-438.
- Over, D. E., & Cruz, N. (in press). Probabilistic accounts of conditional reasoning. In L. J. Ball, & V. A. Thompson (Eds.), *International handbook of thinking and reasoning*. Hove, UK: Psychology Press.
- Politzer, G., & Baratgin, J. (2016). Deductive schemas with uncertain premises using qualitative probability expressions. *Thinking & Reasoning*, 22(1), 78-98.
- Ramsey, F. P. (1990). General propositions and causality. In D. H. Mellor (Ed.), *Philosophical papers* (pp. 145-163). Cambridge, UK: Cambridge University Press. (Original work published 1929).
- Rips, L. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, 90, 38-71.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publications.
- Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98-112). Oxford, UK: Blackwell.