

Miguel López Astorga
Talca University (Chile)

EVOLVED MECHANISMS VERSUS UNDERLYING CONDITIONAL RELATIONS

Abstract. The social contracts theory claims that, in social exchange circumstances, human reasoning is not necessarily led by logic, but by certain evolved mental mechanisms that are useful for catching offenders. An emblematic experiment carried out with the intention to prove this thesis is the first experiment described by Fiddick, Cosmides, and Tooby in their paper of 2000. López Astorga has questioned that experiment claiming that its results depend on an underlying conditional logical form not taken into account by Fiddick, Cosmides, and Tooby. In this paper, I propose an explanation alternative to that of López Astorga, which does not depend on logical forms and is based on the mental models theory. Thus, I conclude that this other alternative explanation is one more proof that the experiment in question does not demonstrate the fundamental thesis of the social contracts theory.

Keywords: conditional relation, evolved mechanisms, mental models, reasoning, social contracts.

Introduction

The social contracts theory (Cosmides, 1989; Cosmides & Tooby, 1992; Fiddick, 2004; Fiddick, Cosmides & Tooby, 2000; Fiddick & Erlich, 2010; Fiddick, Spampinato & Grafman, 2005; Gigerenzer & Hug, 1992) holds that nature, by means of evolution, has provided human beings certain adapted algorithms or mechanisms that are absolutely necessary for their survival. These algorithms or mechanisms help them detect individuals that do not fulfill social deals. Thus, because the fulfillment of agreements is important for the preservation of people and societies, it can be said that such algorithms or mechanisms are essential for human life.

The adherents of the social contracts theory have carried out many experiments in order to demonstrate their theses. Usually, the idea is to show that there are situations related to social agreements in which individuals draw conclusions and in which logic does not appear to play a role.

A representative experiment in this way is Fiddick et al.'s (2000) first experiment.

However, López Astorga (2010) has questioned and criticized that experiment. In his view, there is an underlying conditional relation in the condition of Fiddick et al.'s (2000) first experiment in which it is not possible to observe an explicit logical form. López Astorga's (2010) argument is that, although the logical form is not immediately obvious in that condition, the participants can note an implicit conditional relation linked to the logical form $p \rightarrow q$. López Astorga bases his statement on Almor and Sloman's (2000) study. According to Almor and Sloman (2000), when individuals face an incoherent text, they alter the forms of the expressions included in them in order to obtain consistency.

However, López Astorga's (2010) account has two problems. Firstly, it does not describe in detail the exact process that leads one to discover the conditional relation hidden in the corresponding experimental condition of Fiddick et al.'s (2000) first experiment. Secondly, it does not explain why people tend to execute incorrectly reasoning tasks similar to that of the mentioned experiment when such tasks do not refer to situations of social exchange. Nevertheless, those problems can be overcome by means of a different explanation based on the semantic approach of the mental models theory (Byrne & Johnson-Laird, 2009; Johnson-Laird, 1983, 2001, 2006, 2010, 2012; Johnson-Laird & Byrne, 2002; Johnson-Laird, Byrne, & Girotto, 2009; Khemlani & Johnson-Laird, 2009; Oakhill & Garnham, 1996; Orenes & Johnson-Laird, 2012). This theory can also account for the results of Fiddick et al.'s (2000) first experiment without supposing evolved mechanisms or adaptive algorithms, and this fact can be interpreted as one more proof that those results do not necessarily prove the theses of the social contracts theory. But, in addition, as said and I will try to show below, the explanation based on the mental models theory do not have the two problems of López Astorga's (2010) account.

The mental models theory states that human beings reason paying attention to the possibilities or models related to propositions, their contexts, and their meanings. In this way, the explanation of the human reasoning processes given by that theory enables to understand how, in experimental conditions such as those that will be analyzed below, individuals can note underlying relations (in this case, conditional relations). To show this, I will begin by commenting in more detail on Fiddick et al.'s (2000) first experiment, their results, and the interpretation of them proposed by Fiddick et al. Then, I will explain López Astorga's (2010) criticism of that experiment and the weaknesses of his explanation. Finally, I will show how

the mental models theory can overcome those weaknesses and, therefore, strengthen López Astorga's (2010) idea that, in some occasions, although conditional relations are not explicit, they really exist.

The social contracts theory: adapted mechanisms versus logical relations

As said, the social contracts theory raises that, in social contract situations, i.e., in social exchange situations, logic does not necessarily lead human reasoning. Fiddick et al. (2000) try to prove this by means of an interesting experiment: their first experiment.

That experiment has two conditions: the Conditional version and the Want version. Both of them are based on the Wason selection task (Wason, 1966, 1968). In this way, the Conditional version tells the story of a farmer that has a large number of potatoes and needs to sell some of them. For this reason, he travels to another village in which his language is not spoken. But the fact is that he talks to four individuals living in the village and can understand that they tell him:

If you give me some potatoes, then I will give you some corn (Fiddick et al., 2000, p. 28).

The participants can see four cards, but only one side of each of them. Each card stands for an individual from the village and indicates, on one of its faces, what the farmer gives him and, on its other side, what he gives the farmer. Thus, the first card (from now on, the p card) shows "You gave this person potatoes" (Fiddick et al., 2000, p. 28), the second card (from now on, the $\neg p$ card) shows "You gave this person nothing" (Fiddick et al., 2000, p. 28), the third card (from now on, the q card) shows "This person gave you corn" (Fiddick et al., 2000, p. 28), and the fourth card (from now on, the $\neg q$ card) shows "This person gave you nothing" (Fiddick et al., 2000, p. 28). So, participants' task is to select the card(s) that need(s) to be turned in order to check whether or not the previous conditional rule has always been respected.

Most participants in this condition selected the p and $\neg q$ cards, which, obviously, is the correct answer according to the social contracts theory. Certainly, by choosing p and $\neg q$, the participants can detect whether somebody has broken the rule. The p card allows knowing whether the individual that received some potatoes from the farmer gave him some corn,

and the $\neg q$ card allows knowing whether the individual that did not give the farmer any corn received some potatoes from him. As far as the $\neg p$ and q cards are concerned, it can be said that they are not relevant. To select $\neg p$ is not necessary because, if the farmer did not give that individual any potatoes, it is not possible that that same individual has deceived him. If the farmer did not give any potatoes, he does not expect to receive any corn. On the other hand, to choose the q card is not necessary because, if an individual gave some corn, evidently, that same individual cannot have deceived the farmer.

Thus, it can be thought that these results demonstrate that the social contracts theory holds and that human beings have adaptive mental mechanisms that allow them to note situations in which an individual cheats. However, the problem for the social contracts theory is that these same results are also consistent with logical systems such as propositional calculus. The logical structure of the conditional rule is $p \rightarrow q$, and, if this fact is taken into account, the following explanation is also acceptable:

- If p is selected, it can be assumed that the participant considers $p \rightarrow q$ and p to be premises, and that he (or she) wants to check whether or not the conclusion is q . As is well-known, to do this is to apply the rule of *Modus Ponens*.
- If $\neg q$ is chosen, it can be thought that the participant considers $p \rightarrow q$ and $\neg q$ to be premises and that he (or she) wishes to check whether or not the conclusion is $\neg p$. As is also well-known, to do this is to use the rule of *Modus Tollens*.
- If $\neg p$ is elected, it can be said that the participant considers $p \rightarrow q$ and $\neg p$ to be premises and that he (or she) tries to check whether or not the conclusion is $\neg q$. Nevertheless, as is well-known too, to do this is to make the denying the antecedent fallacy.
- Finally, if q is selected, it can be assumed that the participant considers $p \rightarrow q$ and q to be premises and that he (or she) wishes to check whether or not the conclusion is p . Nonetheless, as is also well-known, to do this is to make the affirming the consequent fallacy.

Therefore, to choose p and $\neg q$ is to use two correct logical rules, and to select $\neg p$ and q is to make two logical fallacies. It can be interpreted that this account implies that the Conditional version raised by Fiddick et al. (2000) is not strong enough to prove the theses of the social contracts theory, since its results can also be explained from a framework based on formal logic. However, the proponents of the social contracts theory have at least two arguments against this last idea. The first one is the fact that people do not often select the correct cards in other versions of the Wason selection

task without social contracts, especially, though not solely, in the initial abstract versions of that task. In such versions, the rule usually indicates that, if a card shows a vowel, then it must have an even number on its other face, and what appears in each card is the following:

- p : a vowel.
- $\neg p$: a consonant.
- q : an even number.
- $\neg q$: an odd number.

Most individuals that execute versions similar to this one choose the cards p and q , which is, as can be checked, a logically invalid answer.

The second argument is provided by the Want version of Fiddick et al.'s (2000) first experiment. The goal of that experimental condition appears to be to show that the expressions in natural language referring to conditional relations (especially, the terms 'if' and 'then') are not really needed. According to them, individuals can solve these tasks even if they do not include expressions related to logical structures. And this is so because individuals do not execute these reasoning exercises following logical rules. What happens is that such tasks trigger the action of the corresponding adapted mental mechanisms for identifying offenders. In this way, in the Want condition, although the basic scenario is maintained, the conditional rule is removed and, instead of it, this text is added:

'I want some potatoes.' You, in turn, know a little bit of their dialect, and tell them **'I want some corn'** (Fiddick et al., 2000, p. 28, bolds in text).

This is the only difference between the Conditional version and the Want version. Everything else is identical in the two versions. But what is important to this paper is that the participants continue to select the same cards, i. e., p and $\neg q$, in the Want version. Given that there is no conditional rule (and hence it is very hard to detect a conditional relation) in this last version, Fiddick et al. (2000) interpret these results as an evident proof that the correct cards are selected due to the use of evolved mechanisms, and that logic does not play a role here. In their opinion, if there is no proposition such as $p \rightarrow q$, the rules of *Modus Ponens* and *Modus Tollens* cannot be applied and the denying the antecedent and affirming the consequent fallacies cannot be made. In short, the explanation cannot be simply logical. However, as indicated above, López Astorga (2010) questions these arguments.

Underlying logical forms

Really, López Astorga's (2010) criticism of the social contracts theory is not the only criticism directed against this theory. Many more criticisms can be found in the literature on cognitive science. Those of Beller (2010), Beller and Spada (2003), or Girotto and Tentori (2008) can be only some examples. There are even theories, such as that of deontic logic (Cheng & Holyoak, 1985, 1989; Fodor, 2000), that, from other perspectives and suppositions, try to explain the same phenomena and results. Obviously, the mental models theory can also be considered to be another alternative approach. In addition, López Astorga has also questioned the social contracts theory in other papers (e.g., López Astorga, 2011, 2013a, 2013b). Nevertheless, in this section, I will only focus on López Astorga's (2010) criticism of Fiddick et al.'s (2000) first experiment. As said, that experiment is one of the most representative of the social contracts theory, and, therefore, it is worthwhile reviewing in detail López Astorga's criticism of it.

López Astorga (2010) does not consider the conditional version to be a problem. In that experimental condition, nothing needs to be explained. As shown above, Fiddick et al.'s (2000) participants gave logically valid answers, which can be understood as simply taking two logical rules (*Modus Ponens* and *Modus Tollens*) and two well-known fallacies related to the conditional (affirming the consequent and denying the antecedent) into account.

The difficulties are caused by the Want version. In this version, in principle, no conditional relation can be observed, and hence an explanation based on the rules of *Modus Ponens* and *Modus Tollens* and the fallacies of affirming the consequent and denying the antecedent cannot be proposed. However, as indicated, López Astorga (2010) resorts to Almor and Sloman's (2000) study. Almor and Sloman used texts taken from Gigerenzer and Hug (1992) and made incoherent reasoning tasks based on them. Basically, the incoherence consisted of using, instead of conditional rules with the structure $p \rightarrow q$, conditional rules such as $q \rightarrow p$. So, although, given the scenarios described in the texts, the appropriate logical form was $p \rightarrow q$, they used propositions inconsistent with the stories in which the antecedent (p) appeared as consequent (q) and the consequent (q) appeared as antecedent (p).

The most important finding, at least for this paper, obtained by Almor and Sloman (2000) is that they noted that their participants did not process the conditional rules as they literally appeared. To the contrary, their participants tended to transform the conditional rules into coherent propositions. Thus, when they deemed it necessary, they transformed propositions

such as $q \rightarrow p$ into propositions such as $p \rightarrow q$. Almor and Sloman (2000) could check this fact because, when their participants completed their tasks, they asked them to remember the conditionals rules. The surprise was that their participants did not remember the verbatim rules ($q \rightarrow p$), but rules transformed by themselves and coherent with the scenarios ($p \rightarrow q$).

From these data, Almor and Sloman (2000) thought that reasoning has two phases. Firstly, mental representations are made and, only after that, reasoning really works. In this way, based on those arguments, López Astorga (2010) states that, if we assume that it is true that those two phases are involved in reasoning, the Want version can be explained in a logical way. That version presented two separated statements: “I want some potatoes” (p) and “I want some corn” (q), but, in their first phase, i.e., in the phase in which the mental representation is made and individuals try to interpret information in a coherent and consistent way, Fiddick et al.’s (2000) participants could make mental representations that somehow linked p and q . Evidently, given the scenario of Fiddick et al.’s (2000) first experiment, the more adequate link between p and q seems to be a conditional relation ($p \rightarrow q$), and, for this reason, López Astorga (2010) claims that this experiment does not prove that human beings have evolved or adapted mental mechanisms for detecting offenders.

So, it can be said that López Astorga’s (2010) idea is that, although p and q are different propositions that are not related in the Want version, Fiddick et al.’s (2000) participants made a logical formalization in which they related such propositions by means of material conditional. Undoubtedly, this idea is interesting, but it must be completed, since the detailed process that leads from the text used by Fiddick et al. (2000) in their Want version to the logical form $p \rightarrow q$ still needs to be described more clearly. In addition, it does not explain why individuals do not offer an adequate answer in versions of the Wason selection task such as, for example, the abstract version previously commented on, i.e., in versions of the Wason selection task without social contracts. As mentioned above, the mental models theory not only can explain the results of the Want version, but also solve these difficulties.

Mental models and logical relations

The mental models theory is a semantic framework. According to it, syntax is not important to human reasoning. Individuals reason considering the models or possibilities that can be consistent in a particular scenario or

situation. A complete exposition of the essential theses of this theory would be very extensive. For this reason, and because that complete exposition is not really necessary if the aims of this paper are taken into account, I will only consider the aspects of the mental models theory directly related to Fiddick et al.'s (2000) first experiment, i. e., the aspects of the mental models theory related to conditional reasoning.

The mental models theory states that three fully explicit models correspond to the conditional. Those models refer to the three combinations or cases in which, according to its true table, a conditional is true. Therefore, the models of the conditional are:

- (A) $p \wedge q$
- (B) $\neg p \wedge q$
- (C) $\neg p \wedge \neg q$

As can be noted, the combination lacking is $p \wedge \neg q$, i. e., the combination that makes the conditional false. This is obvious if we remember that $v(p \rightarrow q) = 0$ only if $v(p) = 1$ and $v(q) = 0$. But the most important point here is that, based on this, the results of the Conditional version can also be explained. The participants mostly tended to select the cards p and $\neg q$, i. e., the cards corresponding to the situation in which the rule is false. So, from this approach, the Conditional version continues without being a problem.

On the other hand, as is well-known, the mental models theory can also easily account for the majority response in the abstract versions of the Wason selection task. As mentioned, (A), (B), and (C) are the fully explicit models, but people cannot always identify all of them. In many occasions, they only can pay attention to the mental models of a proposition, that is, to the models that do not require much effort to be detected. In the case of conditional, the only evident mental model is (A), and this fact explains why individuals tend to elect the cards p and q in the abstract versions of the Wason selection task. Such versions are very complex and hard, and, because they are also too abstract, it can be difficult for the participant to note that the rule refers to (B) and (C) as well.

As far as the Want version is concerned, it can be said that, given that it does not have an explicit logical relation, that circumstance leads us to take all the possible combinations into account and to add that lacking in the previous case:

- (D) $p \wedge \neg q$

But it is not hard to note that, although there is no clear logical relation in this experimental condition, (A), (B), and (C) continue to be valid situations. If the farmer gives some potatoes and the other individual gives

some corn (A), the deal has been fulfilled. On the other hand, if the farmer does not give any potatoes and the other individual gives some corn (B), the farmer has not been deceived. Finally, if the farmer does not give any potatoes and the other individual does not give any corn (C), the situation continues to be right. The only problem is, again, that the farmer gives some potatoes and the other individual does not give any corn (D). This means that the appropriate models in the Want version are (A), (B), and (C) as well. Thus, given that these last three models are the models corresponding to conditional, it can be said that the logical relation underlying the Want version is also a conditional relation between p and q , and that, in this way, a semantic analysis of possibilities is consistent with the selection of the p and $\neg q$ cards too.

The mental models theory hence also explains the results of the Want version. Its advantage over López Astorga's (2010) arguments is that it describes the exact process through which the participants recovered the implicit conditional relation. Nevertheless, the explanation of the mental models theory is not necessarily incompatible with that of López Astorga (2010). Although the mental models theory is based on semantics, it is evident that the acceptance of the models (A), (B), and (C), and the rejection of the model (D) result in a logical relation that can be formalized as $p \rightarrow q$. Likewise, the first phase described by Almor and Sloman (2000) can perfectly be a phase in which the semantic possibilities of propositions are analyzed.

Undoubtedly, we can ignore syntax and assume that all the cognitive phenomena can be semantically explained by the mental models theory. In fact, there are important reasons to adopt this last theory. Not only does it enable to account for participants' results in most experimental reasoning tasks (including, of course, those using abstract versions and versions with social contracts of the Wason selection task), but also it can predict such results. Likewise, as far as I know, it is the only theory that can explain and predict phenomena such as those related to illusory inferences (see, e.g., Khemlani & Johnson-Laird, 2009) in a simple and easy way.

Nevertheless, in my view, to assume the mental models theory in entirety can be problematic. Its main difficulty is that it eliminates syntax in human reasoning and, as can be checked, for example, in Braine and O'Brien (1998), it seems that we have empirical evidence that human reason follows certain 'syntax of thought', that individuals speaking different languages understand some basic logical relations, and that several rules or schemata – e.g., Chrysippus' modus ponendo tollens [$\neg(p \wedge q)$, p ; ergo $\neg q$] and modus tollendo ponens ($p \vee q$, $\neg p$; ergo q) – are naturally used by people.

Therefore, the problem appears to be to find the link between semantics and syntax in human reasoning. Perhaps the right way to go can be that based on O'Brien's (1998) idea about a mental logic. According to O'Brien (1998), the idea of a mental logic in the human mind is not exclusive. So, it can be thought that the human mind can resort to both logical rules and mental models. Maybe this last hypothesis is the hypothesis that needs further exploration, since, if we use syntactic rules and mental models are necessary in cases in which, as in the Want version of Fiddick et al.'s first experiment, there are no obvious logical forms, the only solution seems to be to accept that human inferential activity needs both semantics and syntax.

Conclusions

If the previous arguments are right, a fact is clear: the social contracts theory is questionable. Regardless of other criticisms that can be found in the literature of cognitive science (some of them have been mentioned above), López Astorga's (2010) arguments, if complemented with the semantic framework of the mental models theory, can be considered to be an alternative solid explanation of the same phenomena that the social contracts theory predicts. In fact, as shown, the mental models theory alone can be an alternative explanation. So, if it is true that the human mind has evolved mental mechanisms for identifying cheaters, the proponents of the social contracts theory need further evidence to demonstrate it.

Certainly, the discussion is still open. Of course, it is possible that, in social exchange situations, the logical forms are not relevant and that individuals draw conclusions by means of adapted algorithms designed to detect people that do not fulfill deals or agreements. However, as said, while there are alternative explanations of the experimental results obtained by the social contracts theory, we cannot be sure of that. In this connection, it cannot be forgotten that the mental models theory not only can explain such experimental results. As said, the mental models theory, has a wide scope and can describe, explain, and predict many cognitive phenomena related to human reasoning. Therefore, it seems convenient to take seriously the account presented in the preceding pages.

In any case, there is a fact in favor of the social contracts theory that can be highlighted. Fiddick et al.'s (2000) first experiment is an emblematic and crucial experiment of this theory, but not the only one. The adherents of the social contracts theory have carried out many more experiments in order to prove their theses, and it is obvious that, if we want to come to

a definitive conclusion on this issue, what is appropriate is to review such experiments (or most of them) and to check whether or not their results are also consistent with an explanation similar to the account described in this paper. Nevertheless, because many of those experiments have similar structure, it is probable that the explanation of this paper can be generalized and that, with minor modifications, it can be considered to be coherent with participants' execution in other experiments of the social contracts theory.

However, maybe the most important point is that which refers to the relation between syntactic approaches such as that of López Astorga and the semantic framework of the mental models theory. As mentioned, it can be thought that both accounts are not necessarily incompatible, that the mental models reveal the real logical forms of propositions, and that they only play a role in Almor and Sloman's (2000) first phase. Nevertheless, this idea is also problematic. As argued by Johnson-Laird (2010), the mental models theory can describe and predict human reasoning without resorting to logical forms (the account in accordance with this theory of the Want version of the experiment analyzed in this paper proves this fact too), and it can be said the mental activities involved in reasoning based on syntactic forms and in reasoning based on semantic models are very different. Nonetheless, given that it seems that some links between syntax and semantics can be found in human reasoning, it can be worth continuing to research in this direction.

Acknowledgments

This paper is a result of Project N. I003011, "Algoritmos adaptativos e inferencias lógicas con enunciados condicionales", supported by the Directorate for Research of Talca University, Chile (Dirección de Investigación de la Universidad de Talca, Chile). The author and lead researcher of this Project would like to thank the mentioned institutions for their help in funding this paper and the anonymous expert that reviewed it for his (or her) comments, which, undoubtedly, improved its final version.

R E F E R E N C E S

- Almor, A., & Sloman, S. A. (2000). Reasoning versus text processing in the Wason selection task: A non-deontic perspective on perspective effects. *Memory & Cognition*, *28*, 1060–1069.
- Beller, S. (2010). Deontic reasoning reviewed: Psychological questions, empirical findings, and current theories. *Cognitive Processing*, *11*, 123–132.

- Beller, S., & Spada, H. (2003). The logic of content effects in propositional reasoning: The case of conditional reasoning with a point of view. *Thinking & Reasoning*, *9*(4), 335–378.
- Braine, M. D. S., & O'Brien, D. P. (Eds.) (1998). *Mental logic*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.
- Byrne, R. M. J., & Johnson-Laird, P. N. (2009). “If” and the problems of conditional reasoning. *Trends in Cognitive Science*, *13*, 282–287.
- Cheng, P., & Holyoak, K. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.
- Cheng, P., & Holyoak, K. (1989). On the natural selection of reasoning theories. *Cognition*, *33*, 285–313.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, *31*, 187–276.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York: Oxford University Press.
- Fiddick, L. (2004). Domains of deontic reasoning: Resolving the discrepancy between the cognitive and moral reasoning literatures. *Quarterly Journal of Experimental Psychology*, *57A*, 447–474.
- 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.
- Fiddick, L., & Erlich, N. (2010). Giving it all away: Altruism and answers to the Wason selection task. *Evolution and Human Behavior*, *31*, 131–140.
- Fiddick, L., Spampinato, M. V., & Grafman, J. (2005). Social contracts and precautions active different neurological systems: An fMRI investigation of deontic reasoning. *NeuroImage*, *28*, 778–786.
- Fodor, J. (2000). Why we are so good at catching cheaters. *Cognition*, *75*, 29–32.
- Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating and perspective change. *Cognition*, *43*, 127–171.
- Giroto, V., & Tentori, K. (2008). Is domain-general thinking a domain-specific adaptation? *Mind & Society*, *7*(2), 167–175.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science on language, inference and consciousness*. Cambridge, MA: Harvard University Press.
- Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in Cognitive Science*, *5*, 434–442.
- Johnson-Laird, P. N. (2006). *How we reason*. Oxford, UK: Oxford University Press.
- Johnson-Laird, P. N. (2010). Against logical form. *Psychologica Belgica*, *5*(3&4), 193–221.

- Johnson-Laird, P. N. (2012). Inference with mental models. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 134–145). New York: Oxford University Press.
- Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.
- Johnson-Laird, P. N., Byrne, R. M. J., & Girotto, V. (2009). The mental models theory of conditionals: A reply to Guy Politzer. *Topoi*, *28*, 75–80.
- Khémilani, S., & Johnson-Laird, P. N. (2009). Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, *37*(5), 615–623.
- López Astorga, M. (2010). Los contratos sociales: ¿Un nuevo formalismo ético? *Alpha*, *30*, 9–26.
- López Astorga, M. (2011). ¿Existen reglas adaptativas en la mente humana? Una revisión crítica de las teorías cognitivas evolutivas por medio del enfoque de la perfección del condicional. *Cinta de Moebio*, *41*, 123–143.
- López Astorga, M. (2013a). Lógica deóntica y algoritmos adaptativos: La relevancia del beneficio explícito en las versiones de la tarea de selección. *Filosofía Unisinos*, *14*(1), 52–69.
- López Astorga, M. (2013b). Alcoholism and conditional reasoning: Difficulties in specific mental domains or in general use of heuristics? *Psychologica Belgica*, *53*(4), 3–16.
- Oakhill, J., & Garnham, A. (Eds.) (1996). *Mental models in cognitive science: Essays in honour of Phil Johnson-Laird*. Hove, UK: Psychology Press.
- O'Brien, D. P. (1998). Mental logic and irrationality: We can put a man on the moon so why can't we solve those logical reasoning problems? In M. D. S Braine & D. P. O'Brien (Eds.), *Mental logic* (pp. 23–43). Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.
- Orenes, I., & Johnson-Laird, P. N. (2012). Logic, models, and paradoxical inferences. *Mind & Language*, *27*(4), 357–377.
- Wason, P. C. (1966). Reasoning. In B. Foss (Comp.), *New horizons in psychology* (pp. 135–151). Harmondsworth (Middlesex), UK: Penguin.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*, 273–281.