

Critical Review

Mercier and Sperber's Argumentative Theory of Reasoning: From the Psychology of Reasoning to Argumentation Studies

CRISTIÁN SANTIBÁÑEZ YÁÑEZ

*Centro de Estudios de la Argumentación y el Razonamiento
Facultad de Psicología
Universidad Diego Portales
Santiago, Chile
56-2-6762539*

CRISTIAN.SANTIBANEZ@UDP.CL

Abstract: Mercier and Sperber (2011a, 2011b; Mercier, 2011a, 2011b, 2011c, and 2011d) have presented a stimulating and provocative new theory of reasoning: the argumentative theory of reasoning. They maintain that argumentation is a meta-representational module. In their evolutionary view of argumentation, the function of this module would be to regulate the flow of information between interlocutors through persuasiveness on the side of the communicator and epistemic vigilance on the side of the audience. The aim of this paper is to discuss the perspective of the authors in which they conceive this competence as the natural scenario of reflective reasoning.

Résumé: Mercier et Sperber (2011a, 2011B; Mercier, 2011a, 2011B, 2011c, 2011d) ont présenté une nouvelle théorie stimulante et provocatrice de raisonnement: la théorie argumentative de raisonnement. Ils soutiennent que l'argumentation est un module de méta-représentation. Selon leur point de vue évolutionnaire de l'argumentation, la fonction de ce module serait d'ajuster le flux d'informations entre les interlocuteurs par le biais de persuasion sur le côté du communicateur et par le biais de la vigilance épistémique sur le côté de l'auditoire. Le but de cet article est de discuter de la perspective des auteurs selon laquelle ils conçoivent cette compétence comme le scénario naturel du raisonnement réfléchissant.

Keywords: argumentation, cognition, confirmation bias, evolution, psychology of reasoning, relevance.

1. Introduction

According to Hugo Mercier and Dan Sperber (M&S),¹ their book titled *The Argumentative Theory* is under contract to appear in 2012. For the field of the study of argumentation, this is very good news. A new effort to explain both one of the most complex individual competences and one of the most crucial social practices is always very welcome. And it is even more special when one of the authors (Sperber) has a long career in research and has challenged positions in various fields of the humanities, social sciences and science in general, and the other (Mercier) has a promising career judging by the number of publications on the issue that the author has written on his own or in collaboration with others in a very short time.

But, the good news ends here for those who take the standard approaches in the realm of argumentation studies, namely those who reflect on informal logic, pragma-dialectics, the new rhetoric, dialectics, logic, and usually publish in *Informal Logic*, *Argumentation*,² *Philosophy & Rhetoric*, just to mention a few journals. In what the authors call the target article “Why do humans reason? Arguments for an argumentative theory” (2011a)—the paper in which probably the main ideas of the future book are outlined—M&S reverse the way in which argumentation is usually addressed. In fact, the mere title of the target article is not a question scholars often investigate or ask themselves in the standard study of argumentation.

The authors claim that while the question of *how* humans reason has been investigated thoroughly in many areas (from philosophy to rhetoric), not much has been said about *why* humans have this faculty, capacity or behaviour.³ The question of *why* in the target article (TA henceforth) has been answered by the authors from an evolutionary, cognitive and psychological perspective. By using some results of the experiments of the last 40 years, principally in decision making and reasoning research, they describe the psychological mechanisms that supposedly underlie the actual practices of individuals engaged in argumentative scenarios, such as confirmation bias, polarization, motivated reasoning, among many other mechanisms mentioned in the different papers published by the authors, separately or together.

¹ This information is available at Mercier's online CV: <http://sites.google.com/site/hugomercier/>

² Perhaps, Mercier already could be considered part of the mainstream of this field of study because he has published two articles in *Argumentation*, see bibliography.

³ From the perspective of evolution theory, Desalles (2007) is perhaps one of the exceptions to this generalization.

The authors reverse the common approach to the phenomenon by entitling their theory “The argumentative theory of reasoning.” For Sperber et al. (2010: 383), argumentation is a product of reasoning, whereas reasoning (seen as reflective inferences) is a mental module whose function is to produce and evaluate reasoning manifested in arguments. Sperber and Mercier (in press) point out:

Reasoning can be defined as the ability to produce and evaluate reasons. It is a costly ability: it involves special metarepresentational capacities found only among humans, it needs practice to reach proficiency, and exerting it is relatively slow and effortful. Reasoning, we argue, evolved because of its contribution to the effectiveness of human communication, enhancing content-based epistemic vigilance and one’s ability to persuade a vigilant audience. The reasons reasoning is primarily about are not solipsistic, they are not for private appreciation, they are arguments used, or at least rehearsed, for persuading others.

In the same evolutionary terms, Mercier and Landemore (forthcoming) add that “[c]ontrary to classical cognitive theories, which only provide one level of explanation, that of mechanistic, or proximal (*sensu* Tinbergen, 1963) explanations, the theory delineated ... is an evolutionary theory of reasoning. It argues that reasoning evolved for a specific function—argumentation—that provides a deeper explanation of reasoning’s observed features and pattern of performance.” In other words, argumentation would be an outcome of reasoning only if it is socially expressed.

What is interesting from a standard approach to argumentation is that this new theory overlaps, for example, with the view taken by pragma-dialecticians of argumentation as a social activity (van Eemeren and Grootendorst, 2004: 1), and also with Toulmin, Rieke and Janik’s (1979: 13) general definition of argumentation as an activity of making claims, “challenging them, backing them up by producing reasons,” this is to say, a dialectical social process. But, apart from this preliminary and basic coincidence, Mercier and Sperber’s angle demystifies many of our assumptions and definitions and it directly challenges some more or less common views—such as Johnson’s (2000: 26-7) list of benefits of practicing argumentation.

Mercier and Sperber’s approach was the subject of a special issue of *Behavioral and Brain Sciences* in 2011, where researchers, mainly from a psychology of reasoning angle, commented on the strong and weak points of their theory. Nonetheless, not much has been said from the perspective of dialectical,

logical or rhetorical considerations. The aim of this critical note is to add some remarks by considering the same insights from the psychology of reasoning and evolutionary domain from which the theory originates to see which aspects of the proposal can be scrutinized, questioned, improved or, even, understood.

In Section 2 of this paper, the dual-process theory of reasoning and cognition in general, one of the main supporters of M&S's theory, is summarized and commented upon to see to which extent this could in fact explain argumentation. In Section 3, M&S's notion of *epistemic vigilance* is analyzed to discuss whether this concept sheds new light on the problem of the role that each part of an argumentative exchange has. In Section 4, I focus on, and disagree with, M&S's claim that confirmation bias is evidence of a lack of critical thinking as a natural individual competence in arguing. In the critical remarks section, a general evaluation is put forward and the main critiques are highlighted. To be fair to M&S's theory, these dimensions of the perspective—namely, dual processes theory of reasoning, epistemic vigilance and confirmation bias, among others—have been selected because, in my view, they cover three core and related areas of M&S's proposal, but also cover the phenomena of arguing both as a practical social behaviour and as a cognitive individual disposition and, ultimately, give us the theoretical directions to be sceptical about this new approach.

2. Argumentation in two minds

In his earliest writing on this issue, Sperber (2001: 410) claimed that the function of argumentation "... is linked to communication rather than to individual cognition. It is to help audiences decide what messages to accept, and to help communicators produce messages that will be accepted. It is an evaluation and persuasion mechanism, not, or at least not directly, a knowledge production mechanism." In their last co-authored publications (Mercier and Sperber, 2011a, [the TA]; Mercier and Sperber, 2011b), argumentation is repeatedly presented as a dimension that does not improve cognitive skills and only as a side-effect provides some gains for individuals. The explanation begins by claiming that human thinking is not a homogeneous process governed by intelligence and limited by passions (see also Berthoz, 2009). As empirical research in cognitive psychology has shown—Sperber and Mercier (in press) discuss it extensively—human thinking is a weak mechanism through which humans have conscious access to thought-processes. Reasoning and intelligence cannot refer to any unified process in the mind and, in fact, reasoning and intelligence are reflected in the coordination

or co-function of many modules (see Sperber and Hirschfeld (2007) for a defence of a modularity view of mind).⁴ Some of these modules can even pertain to one of the two systems of the mind, a division that cognitive psychologists have proposed (see Evans and Frankish (2009) for an up-to-date compilation on the issue).

On the hypothesis of the mind being composed of two systems, the ability to engage in argumentation (reflective inferences or reasoning proper) is, at first glance, a manifestation of system 2. Mercier (2011b, p. 85) points out that “‘reasoning’ will be used to refer only to what is usually called system 2, analytic or rule-based reasoning.” But here the problems with M&S’s theory begin. In the literature on the topic (Frankish and Evans, 2009; Evans, 2010; Stanovich 2011; Stanovich and Toplak, forthcoming) there is no clear agreement on what characterizes each system, how they are related, and which processes and functions are inherent to each. For the sake of this promising notion in psychology of reasoning (Evans, 2010), the researchers tend to speak broadly about system 1 as evolutionarily old, unconscious or preconscious, shared with animals, managing implicit knowledge, automatic, fast, parallel, with high capacity, intuitive, contextualized, pragmatic, associative, and independent of general intelligence; and about system 2 as evolutionarily recent, conscious, distinctive of humans, managing explicit knowledge, controlled, slow, sequential, low capacity, reflective, abstract, logical, rule-based, and linked to general intelligence.

One consequence of placing argumentation under this umbrella (the mind as a dual system-modularity architecture) is that the ability to argue is seen as a meta-representational mechanism that, ultimately, is the result of an intuitive multi-unaware mental process. Mercier and Sperber (2011a) try to explain the steps from intuitive inference to reasoning/argumentation proper by combining many concepts and theories in the fields of cognition and evolution:

A process of inference is a process, the representational output of which necessarily or probabilistically follows from its representational input. The function of an infer-

⁴ Sperber and Hirschfeld (2007: 157) boldly maintain that “[a]ccording to the massive modularity hypothesis (see Carruthers 2003; Cosmides & Tooby 1994; Samuels 1998, 2000; Sperber 1996, 2001), the mind is to a large extent made up of a variety of domains—or task-specific cognitive mechanisms or “modules”. It might seem that massive modularity would imply a level of cognitive rigidity hardly compatible with cultural diversity. We want to argue, on the contrary, that massive modularity properly understood is a crucial component in the explanation of this diversity. For a more balanced account of the role of the modularity hypothesis in mind and language, see Carey (2009).

ential process is to augment and correct the information available to the cognitive system. An evolutionary approach suggests that inferential processes, rather than being based on a single inferential mechanism or constituting a single integrated 'system', are much more likely to be performed by a variety of domain-specific mechanisms, each attuned to the specific demands and affordances of its domain (see, e.g., Barkow, Cosmides, & Tooby, 1992). The inferential processes carried out by these mechanisms are unconscious: they are not mental acts that individuals decide to perform, but processes that take place inside their brain, at a "sub-personal" level (in the sense of Dennett, 1969). People may be aware of having reached a certain conclusion, be aware, that is, of the output of an inferential process, but we claim that they are never aware of the process itself. All inferences carried out by inferential mechanisms are in this sense 'intuitive'. They generate 'intuitive beliefs' that is beliefs held without awareness of reasons to hold them. (Mercier and Sperber, 2011a, p. 58)

They continue by adding that the difference between inference and argument relies on the explicitness of the reasons that support a given conclusion in the case of argument:

Arguments should be sharply distinguished from inferences. An inference is a *process* the output of which is a representation. An argument is a complex *representation*. Both an inference and an argument have what can be called a conclusion, but in the case of an inference, the conclusion is the output of the inference; in the case of an argument, the conclusion is a part—typically the last part—of the representation. The output of an inference can be called a 'conclusion' because what characterizes an inferential process is that its output is justified by its input; the way however in which the input justifies the output is not represented in the output of an intuitive inference. What makes the conclusion of an argument a 'conclusion' (rather than simply a proposition) is the fact that the reasons for drawing this conclusion on the basis of the premises are (at least partially) spelled out. As Gilbert Harman has justly argued (Harman, 1986), it is a common but costly mistake to confuse the causally and temporally related steps of an inference with the logically related steps of an argument. (Mercier and Sperber, 2011a: 58)

But adding the idea that in arguments there is a logical relationship between premises and conclusion, and not only causal and temporal related steps, does not clarify the differences between

inference and argument, because the question remains what the difference is between the “spelling out” (a linguistic and social act) and the “logical” aspect. At the same time, the authors sympathize with various frameworks, and this usually leaves the promoters at the edge of a cliff, for instance when they go back to system 1 of the mind—perhaps to maintain the evolutionary mainstream narrative. One gets the idea that they do so in order to explain the elaboration of a reflective mechanism: arguments are intuitions about the relationship between reasons and conclusions. They assert:

... all arguments must ultimately be grounded in *intuitive* judgments that given conclusions follow from given premises. In other words, we are suggesting that arguments are not the output of a ‘system 2’ mechanism for explicit reasoning, that would be standing apart from, and in symmetrical contrast to, a ‘system 1’ mechanism for intuitive inference. Rather, arguments are the output of one mechanism of intuitive inference among many that delivers intuitions about premise-conclusion relationships. Intuitions about arguments have an evaluative component: Some arguments are seen as strong, others as weak. Moreover there may be competing arguments for opposite conclusions and we may intuitively prefer one to another. These evaluation and preferences are ultimately grounded in intuition. (Mercier and Sperber, 2011a: 59)

What is salient in these quotes is that by framing argumentation in this way, they prepare the ground for the introduction of the core explanation: arguments are complex-reflective inferences only if, at least partially, the representations (reasons) are spelled out, that is to say, when the environment—the audience—requires some kind of refinement of the information presented. They explain:

Here we want to explore the idea that the emergence of reasoning is best understood within the framework of the evolution of human communication. Reasoning allows people to exchange arguments that, on the whole, make communication more reliable and hence more advantageous. The main function of reasoning, we claim, is *argumentative* (Sperber, 2000a, 2001, see also Billig, 1996; Dessalles, 2007; Kuhn, 1992; Perelman & Olbrechts-Tyteca, 1969; Haidt, 2001, and Gibbard, 1990, offer a very similar take on the special case of moral reasoning). (Mercier and Sperber, 2011a: 60)

By “argumentative” they mean that reasoning is situated in a dialogue, in a social context, in which individuals—and only

because of this public condition—ponder reasons, even anticipate scenarios, to affect someone's thoughts or actions (from the point of view of the communicators), and to filter information that could derail us (from the point of view of the audience). The heart of the proposal, once the basic evolutionary intersection between the mind and the social is explained, would be that the principal argumentative profile of reasoning serves human communication purposes. This practice improves or increases “both in quantity and in epistemic quality the information humans are able to share” (Mercier and Sperber, 2011a: 60). But this view differs from important proposals about the function of system 2, including the reflective module (see Evans, 2009, 2010; Frankish, 2009; Saunders and Over, 2009; Stanovich, 2011; Stanovich and Toplak, forthcoming). One of the main problems of S&M's approach is that they do not go deeply into the characterization of system 2, which involved the reflective processes. In the conclusion I will come back to these issues.

Humans are good at arguing, although we rarely, according to the authors (quoting empirical research) engage in higher-order reasoning, this is to say, ponder reasons about reasons. In fact, what this theory challenges is that humans make good decisions, maintaining that we prefer to make decisions we can justify more easily in front of others.⁵ The empirical research repeatedly quoted by the authors allows them to say that reasoning is not a higher form of individual cognition, but is better used in collaborative behaviour, because there it produces better outcomes.⁶

In “Reasoning as a Social Competence,” Sperber and Mercier (in press) summarize many of the aspects commented on in this review so far:

Reasoning, we have argued, is a specialized metarepresentational competence with a primarily social cognitive function. It is both structurally and functionally quite different from intuitive inferential mechanisms that have a primarily individual cognitive function. Collective cognitive performance may be based on the aggregation of individual intuitions or on argumentative interaction, with quite different outcomes.... When argumentation and hence reasoning are at work, they shape the outcomes of

⁵ Mercier and Sperber (2011a, p. 69) point out: “According to this theory, people often make decisions because they can find reasons to support them. These reasons will not favour the best decisions, or decisions that satisfy some criterion of rationality, but decisions that can be easily justified and are less at risk of being criticized.”

⁶ Nevertheless, the authors rapidly point out that “... it should be stressed that the argumentative theory does not predict that groups will always make better decisions, but merely that reasoning should work better in the context of a genuine debate” (Mercier and Sperber, 2011b: 99).

group processes. In many cases, this is for the best—more information is shared, superior arguments are granted more weight. Sometimes, however, reasoning creates a polarization of the group (Sunstein, 2002). This mostly happens when people are forced to debate an issue on which they already agree.

Thus, this approach emphasizes that reasoning is a mental act of constructing or evaluating an argument, which contrasts with ordinary intuitive inference—a process that yields a conclusion without articulating the reason to accept it. The authors argue that the main function of reasoning is social, but it serves the social interests of individuals rather than the collective interests of the group. This is shown by the fact that people produce arguments within a “high degree of mere satisficing”, which is to say that (academic and intellectual contexts aside) people do not look for the best formulation of the best argument possible, but instead they use the first minimally decent argument, and if the argument does not work then a rebuttal or another argument is put forward.

As soon as the Target Paper began to circulate, scholars critically assessed it by putting forward important objections. M&S (2011a) mention that the most irritating objection that they really want to correct is that their theory of reasoning has only rhetorical goals: that reasoning is only designed to find arguments to persuade others. M&S think that, on the contrary, reasoning evolved in part to make people change their minds by presenting them with good reasons to do so. But then the question is: what kind of theoretical explanation could balance the idea that reasoning evolved also with epistemic goals—the inclination to give good reasons—and, at the same time, that people make a minimal effort to put forward good arguments and that audiences, more often than not, accept these decent arguments instead of only the best ones?

3. Epistemic vigilance: The defensive warrior metaphor

M&S believe that the misinterpretation of their theory (Desalles, 2011; Evans, 2011) is due to the mistaken overemphasis by the critics on the role of communicators—the role of the communicators being to produce arguments to persuade—instead of seeing the role of the audience, which is to evaluate arguments to choose those that yield useful information.

If reasoning as an argumentative practice is performed by individuals by means of minimal cognitive efforts because it is a high-cost mental activity with a relatively high failure rate, then

why does reasoning as an argumentative practice exist at all? The answer of M&S is:

Humans are immersed in a flow of socially transmitted information and are highly dependent on it. For communication to have evolved, it had to be advantageous to both communicators and receivers.... What makes communications advantageous to receivers is that it provides them with rich information that they could not, or not easily, have obtained on their own. For this, the information they receive has to be genuine information, that is, close enough to the truth. What makes communication adventurous to communicators is that it allows them to achieve some desirable effect in the receivers. For this, the information they emit has to be conducive to this effect, whether it is true or false. (Mercier and Sperber, 2011b: 96)

Because receivers have to be alert in order to avoid misleading information, they must exercise what Sperber et al. (2010) call “epistemic vigilance.” This is not a slogan of the “War on Terror,” but the concept the authors use to stress the active role of the audience in the practice of argumentation. Epistemic vigilance, apparently, is a cognitive skill to filter the information carried out by three heuristic mechanisms or strategies: assessing the trustworthiness of the communicators, checking the coherence of the message and assessing the relevance of the message (although relevance seems to be, for Sperber et al. (2010: 376), only a step in the coherence-checking process).

By assessing the trustworthiness of the communicators, the authors simply mean that receivers automatically calibrate the level of trust of the source of the message. By checking coherence, they mean checking the degree of believability of the information received. The highest degree is obtained, in turn, by the audience through pressure on the communicators by asking them to display more adequate arguments for the discussion at hand. In Sperber's (2001) evolutionary terms, the idea is clarified in the following narrative:

My first suggestion is this: coherence checking—which involves metarepresentational attention to logical and evidential relationships between representations—evolved as a means of reaping the benefits of communication while limiting its costs. It originated as a defense against the risks of deception. This, however, was just the first step in an evolutionary arms race between communicators and audiences (who are of course the same people, but playing—and relying more or less on—two different roles). (Sperber, 2001: 409)

Obviously, when trust is doubtful, communicators lay more stress on the contents of the messages: they commit themselves to utter more coherence-based reasons for the acceptance of a given message. At the same time, smart audiences should balance the reliability of the source with the believability of the content.

Relevance (Sperber and Wilson, 1995) is an omnipresent mechanism through which certain deductive protocols are activated to interpret the message in a more productive way. This does not mean that hearers tend to search for a charitable interpretation; relevance is more or less an asymmetrically proportional measure: when the costs of searching for the right interpretation are too high, then the cognitive effects tend to go down and, for this reason, hearers could abort the mission of searching. In argumentative scenarios, when a piece of information has little relevance or is directly irrelevant, and thus the presumption of relevance is interrupted, then hearers are pushed towards a sharper epistemic vigilance. In other words, in contrast to the basic assumption among argumentation scholars, this idea goes against the principle of charity (for a comprehensive account of this principle in argumentation theory, see Govier, 1987).⁷ The principle of charity states that it would be unfair to arguers to give their discourse anything less than a reasonably sympathetic interpretation, because, in the absence of indications to the contrary, it is presumed that arguers intend to be relevant and reasonable. For Sperber et al. (2010), this behaviour is unnatural: unless we are in very specific settings—teaching children, in critical moments, or in extreme situations—people tend to avoid high-cost interpretative reconstructions.

Reasoning is a tool for epistemic vigilance. It is the “verbalization” (with special constructions such as “if... then...,” “therefore,” etc.) of a sophisticated defence that filters unclear, incomplete, tricky, misleading, or even abusive discursive communication and information. Because plain cooperation, for Sperber et al. (2010), is not necessarily the norm among members of groups, individuals had, from an evolutionary point of view, to develop a kind of epistemic protection.

If we follow the metaphoric-semantic consistency of the notion *epistemic vigilance*, one could ask: to what extent is this protection or vigilance manifested in “preventive attacks”? The authors do not say. In other words: in argumentative scenarios, to ask for clarification, to counter-argue, or to put forward doubts are more than passive mechanisms, which the simple

⁷ In this account, Govier (1987, p. 148) defends, nevertheless, a moderate version of charity.

idea of vigilance seems to convey. For this precise reason, normative argumentation theories have developed standards for a reasonable discussion, critical discussion, and so on. And for this very basic reason, rhetoric has shown the ways in which persuasion takes place in important social affairs, when power, interests, games of predominance, and a long list of human inclinations, are part of extremely calculated exchanges. Audiences are not only in a vigilant stance when they hear a discourse; they are already in the discourse itself, due to the way the message is composed. On many occasions, audiences are co-responsible for what the communicators express.

Beyond this “metaphorical” remark on the concept of epistemic vigilance, a critique from a pure cognitive point of view can also be advanced. There is a well established consensus among cognitive researchers (Sterelny, 2003; Godfrey-Smith, 1996) that all organisms have (and develop) a mechanism or so-called detection system, which consists of a control of behaviour to fit in and to manage the information of the environment and, in this way, to preserve their existence in that environment. The detection system allows the organism both to send out the right quality/amount of information, and to process the information important for their survival. Those detection systems never are perfect, they require flexibility of the organism, and the organism has to develop specific strategies to read, or deal with, opaque context information. In our discussion, the organism is our speaker. If a speaker, as a natural tendency, and even as a first reaction after getting the answer of the audience, persists in its confirmation bias error, then how is her detection system working? It is not enough to say that the receivers, audiences, listeners, convincees, are applying their epistemic vigilance to calibrate the flow of information, because what the evolution of cognitive systems shows is that senders apply, in every act, a control over the information that they send out.

4. Confirmation bias

The premise of the argumentative theory of reasoning is that reasoning evolved in such a way that we can argue with others to determine the weight of information for our daily tasks. This activity, in return, makes us good at doing just that: arguing. According to M&S, in real and truly argumentative scenarios, people are good at finding and evaluating arguments, but bad at doing it in artificial settings in which we ask for the resolution of reasoning problems (this latter point being a commonplace among psychologists of reasoning). At the same time, if reasoning is indeed a mechanism to argue collectively, then group per-

formance is more important than individual performances. Groups argue even better than the best individual of the group. Finally, this theory maintains that if reasoning evolved in order to argue *with others*, then we should be biased in our search for arguments: in a discussion, we tend not to use the arguments that rebut our claims. As the literature in the areas of judgment and decision making in psychology terms the heuristic, we proudly stand in this world showing our “confirmation bias.”

As the authors have repeated (Mercier, 2011a; 2011b; 2011c; 2011d; in press; Mercier and Landemore, forthcoming; Mercier and Sperber, 2011a; Sperber and Mercier, in press), the confirmation bias is the most robust and prevalent heuristic disposition in reasoning. Confirmation bias is usually indicated as being responsible for much of our reasoning mischief. What is biased, nonetheless, is the production of arguments, not their evaluation. Surprisingly, for M&S confirmation bias can be seen also as a sort of “division of cognitive labour” in the activity of arguing, because by using this bias it would not be necessary for all those involved in the discussion to laboriously assess the pros and cons of each option under scrutiny. Again, the emphasis on confirmation bias in this theory challenges basic assumptions in standard approaches to argumentation.⁸

In short, in natural settings, when people reason they do not try to produce the best answer to a given rebuttal, but rather try to find confirmatory arguments that maintain their beliefs. Because confirmation bias is not related with the capacity or ability of grasping misleading information, this bias is, more or less, in perfect balance with the need for openness to negative considerations, such as rebuttals. According to M&S, when we are in the position of evaluating arguments, we are guided by the urge to keep valuable information and, for this same reason, we are open to accept those arguments that force us to revise our beliefs. In genuine deliberations, the confirmation bias of each individual is checked and “compensated by the confirmation bias of individuals who defend another position. When no other opinion is present (or expressed, or listened to), people will be disinclined to use reasoning to critically examine the arguments put forward by other discussants, since they share their opinion” (Mercier and Landemore, forthcoming). Sperber and Mercier (in press) add:

⁸ Certainly, scholars who participate, for example in ISSA and OSSA, are very well aware of the role of confirmation bias in argumentation, but, unless my review is incomplete, in the most recent proceedings of both the conferences it is difficult to find a paper in which all these topics are addressed from a cognitive or an evolutionary point of view.

Is the confirmation bias therefore an aspect of reasoning that may be effective from a practical point of view but that makes reasoning epistemically defective? Not really. People are quite able to falsify ideas or hypotheses...*when they disagree with them*. When a hypothesis is presented by someone else, participants are much more likely to look for falsifying evidence (Cowley & Byrne, 2005). When, for instance, people disagree with the conditional statement to be tested in the Wason selection task, a majority is able to pick the cards that can effectively falsify the statement, thereby successfully solving the task (Dawson, Gilovich, & Regan, 2002). Similarly, when people believe that the conclusion of a syllogism is false—if it conflicts with their beliefs for instance—they look for counterexamples, something they fail to do otherwise (Klauer, Musch, & Naumer, 2000)

Confirmation bias, unfortunately, is something that people cannot suppress just like that. Reasoning, in this sense, has evolved in a way that makes its agents solipsistic machines, unless we are in a group. Only this setting would make epistemic benefits possible. Interestingly, Mercier (2011a: 313) remembers that “[i]f argument quality is not sufficiently high in a domain, the confirmation bias will make experts tap into their vast knowledge to defend whatever opinion they hold, with polarization and overconfidence as expected results.” Everyone, perhaps, has experienced discussions with very knowledgeable people who generate good pro arguments but, at the same time, avoid spelling out con arguments. But does this last situation not depend on the strategy of the speaker? We also need an evolutionary and cognitive account of argumentative strategic competence.⁹

5. Critical remarks

For a more traditional researcher in the argumentation arena, certainly many basic dimensions of argumentation are not taken into account in M&S's approach. The list is long. Some important ones are: the problem of argumentation as a commitment-based activity (Walton and Krabbe, 1995; Walton, 2007); the role of language and its pragmatic dimension (here the pragma-dialectician developments may be quoted); the place of formal and informal logic rules/principles (Johnson, 2000); the prob-

⁹ Johnson-Laird (2011) has a kind of proposal on this matter by using the label of “metareasoning” to refer to the notion of “strategic reasoning” in which a strategy “is a sequence of steps that enables us to infer the answer to a deductive problem” (262). Obviously, these ideas are still captive of describing only the isolated inference process, and not the process of proper argumentative scenarios.

lems of practical reasoning (see the critical commentary of Narvaez (2011) on M&S's TA); and, especially, the whole problem of normativity.

It would be unfair to measure M&S's proposal by questioning it only from the point of view of dialectics, rhetoric or logic, considering how distant those areas (the "how" considerations, according to M&S) are from their own focus (on the "why" considerations). Nonetheless, normativity also has been explained from an evolutionary, cognitive and communicative perspective with crucial implications for an argumentative theory of reasoning and rationality in general (Oaksford and Chater, 1998, 2007; Stanovich, 1999; Stein, 1996; Sterelny, 2003; Tomasello, 2010, 2008). But what is more striking for me is that normativity has been one of the most central targets in the domain of psychology of reasoning (for a meta-discussion on the topic, see Elqayam and Evans, 2011), where interesting types of normativity and closely related notions have been distinguished, such as: empirical normativism, prescriptive normativism, empirical logicism, prescriptive logicism, empirical Bayesianism, prescriptive Bayesianism.

Why would those general notions be of interest in describing and explaining reasoning in an argumentative vein? For two reasons. First, they are conceptual constellations in the psychology of reasoning where a complete package of theoretical frameworks is given to conceive human reasoning and, at the same time, they are part of the background from which M&S take their insights, although they differ substantially from M&S's proposal in many respects. And second, because they also have methodological parameters by which the ecology of reasoning is studied to support the conceptual frame elaborated, and from which almost all the experiments quoted by M&S are reported.

Thus, for example, the prescriptive normativity of Johnson-Laird's approach (2011, 2005) emphasizes the idea of reasoning as a mental model and representation-pictorial device by which an agent reasons using content-based frames through which he can imagine the significance of a situation, hypothesize it, and select, most of the time, the common feature of some possibilities of the inferential situation at hand, and in this way resolve, for example, conditional scenarios. According to Johnson-Laird (2011, p. 209 ff), unless we have knowledge to the contrary, we assume that each model represents an equiprobable alternative to deal with a contextual inferential challenge. It is easy to observe the idea of defeasibility argumentation behind this point of Johnson-Laird's approach. However, M&S say nothing about defeasibility, even though this is a core problem that must be addressed in building a solid theory

about argumentation or reasoning in general. M&S do not even criticize the idea of defeasibility in Johnson-Laird.

By contrast, *Bayesian rationality* (Oaksford and Chater, 2007), highly empirical and prescriptive according to Elqayam and Evans (2011), conceptualizes reasoning as a probabilistic *machine* dealing with an uncertain world; a machine, nonetheless, influenced by the conversational pragmatics affecting subjective probabilities. For Oaksford and Chater (2007), the way in which the cognitive system works is grounded in Bayesian probability and normatively justified by it. This means, in short, that the *rational* behaviour of human beings exemplified by everyday reasoning and decision-making is, so far as human evolution demonstrates, a successful action, individually and in groups. This idea clearly is contrary to M&S's message that maintains that individual performances are tied to poor outcomes. M&S also claim, contrary to what Bayesian rationality seems to indicate, that people exhibit confirmation bias all the time without any possibility of choosing alternative arguments. It is quite surprising that while using more or less the same data the two approaches reach opposite conclusions.

Stanovich's dual-system theory characterizes system 1 as heuristic and contextualized, which indeed can be intervened in by the analytic system 2.¹⁰ These interventions, however, are affected by the cognitive ability of the participants. This basic and strongly empirically supported idea does not appear in M&S's angle, namely the idea of a *vivid* exchange between both systems. Is confirmation bias an outcome of system 1 or not? If so, can it be modified by system 2 of the same individual without the intervention of the audience? If M&S think that the latter situation could be the case, they do not explain how. Focusing on the idea of instrumental rationality, Stanovich (2004, 2009) maintains that agents have special abilities to, for example, intervene in the automaticity that system 1 has. Automaticity is composed of two sub-mechanisms at least: emotional regulations and experiential-adaptive associative learning processes. These two sub-mechanisms generate the automatic replies of the agent (or individual) to the input of the environment (Stanovich and Toplak, forthcoming). I adhere to Stanovich and Toplak's (forthcoming) proposal that the main feature of System 2 is the cognitive decoupling. The energy cost of system 2 is not its most important characteristic, nor is the fact that it is serial. Stanovich and Toplak (forthcoming, DOI) differentiate the systems as follows:

¹⁰ Stanovich is recognized as one of the first and main scholars in proposing the notion (Evans, 2010; Kahneman, 2011).

All of the different kinds of Type 1 processing (processes of emotional regulation, Darwinian modules, associative and implicit learning processes) can produce responses that are nonoptimal in a particular context if not overridden. For example, often humans act as cognitive misers (an old theme in cognitive/social psychology...) by engaging in attribute substitution—the substitution of an easy-to-evaluate characteristic for a harder one, even if the easier one is less accurate... For example, the cognitive miser will substitute the less effortful attributes of vividness or affect for the more effortful retrieval of relevant facts... But when we are evaluating important risks—such as the risk of certain activities and environments for our children—we do not want to substitute vividness for careful thought about the situation. In such situation, we want to employ Type 2 override processing to block the attribute substitution of the cognitive miser.

Can we consider argumentative situations, at least some of them, as risky? That is a rhetorical question: clearly this is the case. Can we override our confirmation bias as a first reaction in those dialectical risky contexts? Yes, we can; this is the conviction of Stanovich and Toplak (forthcoming). Perhaps M&S can reply that this exactly is the case because it is the first reaction—the confirmation bias—that is rejected by the audience; therefore it is the audience in its constant epistemic vigilance that is responsible for putting pressure on the agent to make his system 2 override the automatic answer. But then again, if this is the normal process of reasoning/argumentative ecology, why should we emphasize the confirmation bias (the first move of a speaker arguing) as one of the main concepts in a theory about reasoning and argumentation, instead of taking a dialogical perspective in which all the moves are conceptualized? Because M&S focus on system 2 to back their contribution to the dual-system theory on reasoning, reflective behaviour, and, ultimately, human communication, it is odd that the notion of decoupling is not used. This idea is important because it adds to the explanation of system 2, two core processes: suppressing system 1 reactions and activating hypothetical reasoning and cognitive simulation. As Stanovich and Toplak (forthcoming) put it: “In order to reason hypothetically we must, however, have one critical cognitive capability—we must be able to prevent our representation of the real world from becoming confused with representation of imaginary situations. The so-called cognitive decoupling operations are the central features of Type 2 processing that make this possible.” In which part of M&S’s frame of reference could we add the idea of the existence of a critical cognitive capability? As the evidence seems to show (much of the same evidence that

it is quoted by M&S) the ability of sustaining the decoupling process of secondary representation should be part of an argumentative theory of reasoning.

As I said in the beginning of this section, M&S are not the first to try to say something about the function of argumentation in evolutionary terms, that is, to try to answer the question of *why* we argue. A French researcher on cognition, Desalles (2007, p. 294 ff.), explored some conjectures and proposed that the “argumentative mode” (different from the “informative mode”)—which is one of the main driving forces to generate language in our species—arises when agents try to find a solution to a cognitive conflict. Cognitive conflict is the core idea in Desalles. According to Desalles (2007), human beings in argumentative mode detect and resolve cognitive conflict, and a cognitive conflict arises when the beliefs and wishes that constitute the conflict have some degree of contradictory intensity. Human beings have cognitive conflicts when they have different representations of a given situation or of terms by which to frame the situation. Desalles even goes in the direction of presenting a model to understand how human agents resolve cognitive conflicts—a tripartite model: assessment, abduction and negation—but what is extremely surprising is that after explaining the argumentative mode he concludes:

Human beings reason in the same way as they argue; and the assessment-abduction-negation process that they use for reasoning is marked by its conversational origin. A common view of language is that it is an outgrowth of the capacity for reasoning. On this view, human beings, who are intelligent because intelligence is ‘useful’ for their survival, take the advantage of their intelligence to speak and argue. The preceding discussion makes it possible to turn this view on its head and see the capacity for reasoning as an outcome of our argumentative abilities. This would mean that the capacity for reasoning logically derives from the abilities necessary to conversation (Desalles, 2007: 307).

Leaving aside the difference emphasized by M&S between intuitive and reflective reasoning, it is curious that they use the same idea as Desalles’s, namely that language/reasoning evolves because we argue. M&S, however, do not mention anything similar to the notion of “cognitive conflict,” perhaps because this general idea needs to be explained from the point of view of cognitive abilities, considering among them those which make an agent a critical thinker when he argues, something totally unaccepted in M&S’s frame.

So far, I have discussed M&S theory by comparing it, in a general way, with other views within the psychology of reasoning and cognitive approaches, without considering more traditional aspects of argumentation theory proper. I will continue adding other remarks in this vein. By showing that their approach is, in some respects, a compilation of many frameworks and theories, I do not, however, mean to suggest a big failure. For the theoretical task is to use results and insights of other researches to build, in a different accent and light, our own theory. In the argumentation arena, in fact, this is the case with the pragma-dialectics theory.

However, more pressing critical nodes in M&S's approach are the following. Consider, for example, the basic assumption, or hypothesis, of M&S about why reflective reasoning evolved. M&S maintain that the function of reasoning is argumentative, it is to devise and evaluate arguments intended to persuade:

We view the evolution of reasoning as linked to that of human communication. Reasoning, we have argued, enables communicators to produce arguments to convince addressees who would not accept what they say on trust; it enables addressees to evaluate the soundness of these arguments and to accept valuable information that they would be suspicious of otherwise. Thus, thanks to reasoning, human communication is made more reliable and more potent. (M&S, 2011: 72)

So said, reflective reasoning is a collaborative endeavor to improve collective human communication, rather than to provide an individual gain. But, couldn't it be the case that in the flow of evolution of reasoning an exaptation (Gould, 1991) could occur?¹¹ For, granted that groups improve communication of the collective by arguing, yet individuals also reach better decisions and more accurate beliefs when they argue than when they do not. Nonetheless, one could maintain that the evolutionary story could be quite the opposite, with reasoning evolving to contribute to selfish individuals' interests in having more accurate beliefs. In this scenario it helps to be open to cons and accept them when it is convenient to achieve some goals.

At the same time, it is a redundancy to say that we reason better when we argue in groups. Any social activity has better outcomes in social reunions, in groups, even considering the extreme and smallest group: two persons (Brown, 1999). Note that to maintain this kind of claim (e.g., that people reason better in groups), M&S are basing their view on data that is under revi-

¹¹ Wiles (2011) has a similar point when she comments on M&S's TA in the *Behavioral and Brain Sciences* special issue.

sion. There is an ongoing and growing dissatisfaction using the results from artificial settings (Elqayam and Evans, 2011) to understand the confirmation bias. It should be added here that M&S's theory could benefit from adding more details about acting as a group member,¹² for probably this context also has some implications for the way an agent behaves reflexively and argumentatively. It is well established that one main feature of the nature of any group is the exercise of some types of coercion and control. What is the implication of coercion for an argumentative theory of reasoning? Another related issue with the idea of "better in group" is the question whether this characteristic could support the speculation surrounding the idea of reasoning as group selection. M&S (2011: 97) only touch this issue when they respond to one of those who commented on the TA.

When M&S maintain that the function of reasoning is argumentative, that it is to devise and evaluate arguments intended to persuade, they abolish in one shot the distinction between rhetoric and argumentation, whereas the latter field studies precisely the criteria why and how good arguments can finally convince and persuade different audiences. For this reason, some scholars (Godfrey-Smith and Yegnashankaran, 2011; Narvaez, 2011; Poletiek, 2011; Stemberg, 2011) characterize M&S's theory as being a rhetorical one.

Another critical point in M&S's explanation of reasoning as argumentative activity, is that they explicitly emphasize that they are concerned with *why* reasoning evolved, instead of *how* reasoning works, but in order to attain this objective they characterize many *how* issues (such as: confirmation bias; better in groups; the audience is alert and vigilant; the more cultural differences between the individuals discussing, the better the outcome). There is some doubt whether the authors commit the same "fallacious" move of confirmation bias by using and *selecting* the data, results, and insights available in 40 years of research that confirms their approach. One tends to think that there is a paradox in the idea of practicing reflective reasoning, while at the same time producing reactions that are only looking for confirmation.

Nothing is said in M&S's position about the many ways in which someone can argue about, for instance, practical matters. Good politicians (those who win elections arguing with very indirect strategies, or even with the strategy of no arguing at all!) go to meet with the people knowing and saying what the audience wants to hear, confirming the people's values, not confirming the point of view of the speaker. Here I am not emphasizing

¹² For a powerful account of the so-called philosophy of sociality, see Tuomela (2007).

what is a cliché in the internal rhetorician tradition, namely: adjust your speech to the hearer. Instead of remembering this triviality, I am thinking about something that in evolutionary biology, especially concerning human cognition and its social ramifications, is a consensus: the plasticity and adaptability with which the cognitive system adjusts its operation, in a marvellous balance, between its internal structure and the environment. Three questions arise from an evolutionary point of view: does reasoning have plasticity as does its superstructure, the cognitive system? If so, how is it expressed, and, what role does the plasticity of reasoning play in an argumentative scenario? M&S do not offer an answer yet.¹³

The authors open a discussion, in a very nice way, around the possibility of having, theoretically speaking, categories of arguments. Using an analogy the authors ask: Is it possible to have the category “good argument” like, for example, the category “good restaurant”? This reflection of the authors goes against classic argumentation and rhetoric scholars. The authors agree that rhetoric has developed useful classifications in order to reveal strategies, but for them it is not clear in which ways these could help people to look for arguments in daily life. When people want to find just any good restaurant, then the categories “Japanese restaurant,” “Italian restaurant,” or “French restaurant” are rather irrelevant. Depending on the topic, context and interlocutors, any representation could be a good or bad argument (or even not an argument at all). In the case of a restaurant, many things come to your aid, for example the infrastructure of the city in which you are looking for the restaurant (knowing that in some areas there are plenty of good restaurants), which could make your search easier, but M&S ask: are our minds organized in the same helpful way? The answer in their own theory is: no, the mind definitely is not organized in the same way as a city (a context in which you know there will be an unforgettable restaurant!). But M&S forget that many institutional and social scenarios, which are argumentation’s natural settings, are organized in that way and good arguments can be found “encapsulated” in these settings. In these social scenarios, and of course depending on the reaction or answer of your opponent, the information available, the time available to develop the discussion and many other factors, you can find a good argument. The social situation as a whole incorporates the mind of an individual—that individual being one of the resources needed to have a good argument. That is why pragmatodialectics is about a protocol for good argumentations (critical

¹³ See Sterelny (2003) for a detailed account of the evolution of human cognition and its many ways of adaptability and plasticity.

discussion), or why informal logic considers basic criteria to begin with to evaluate a good argument (acceptability, sufficiency, relevance). Good arguments are not in the mind, but in the situation, viewed rhetorically, or in an ideal reasonable way to resolve a dispute, viewed normatively. The critical remark here is then: from the point of view of a theoretical account of the argumentative competence, really good arguments cannot be found in “isolated minds” because the evaluation of such a “products” is tied to information and contexts that collective agents share, both for practical affairs and abstract and artificial settings. Thus, the analogy between mind and city is not appropriate to try to explain where good arguments are located.

Another point of critique, as was discussed earlier, is the notion of “epistemic vigilance.” According to the definition, this is a cognitive “filter” that automatically protects listeners. Reasoning (as reflective inference) was defined as a social device whose main function is argumentative, that is to say, a mechanism that is “turned on” as soon as we open our mouth. But contrarily, “epistemic vigilance” is at work all the time, as a heuristic—fast—device that allows individuals to distinguish between valuable and poor information. One question here could be: is “epistemic vigilance” part of system 1 or system 2? Does “epistemic vigilance” promote individual epistemic benefits instead of collective interests? While leaving these questions open to future research, what reality shows, nonetheless, is that when, for example, we are in mediation or negotiations we often put our, so to say, local epistemic vigilance on “stand-by” in order to obtain a broader goal. In other words, “vigilance” is perhaps a matter of degree and strategy. After all, epistemic vigilance does not need to be verbalized, for we can stay silent forever in a very “sharp vigilant position.” Thus, the main critical observation here is that it is not enough for an argumentative theory of reasoning to emphasise that human beings have developed an epistemic protection. It would be preferable to try to explain of which specific mechanisms this protection is composed of and why human beings under certain circumstances delay, postpone or inhibit that protection.

In my view, another problematic point is the notion of information. Because the angle of the authors is cognitive and psychological about the nature of human communication, information is a core concept. But when we argue, are we really sharing information? All depends, finally, on the notion of information that we have at hand. When discussing whether to walk or take the bus to the cinema, someone says, “It is a beautiful day today!” thereby implying that they should walk. Does the sentence convey information or does the exclamation only convey the communicative intention of the communicator and

nothing of the external world? The authors could reply that this is precisely the information sent: the communicative intention. But then we go into the game of second or higher-order informational processes.

When the authors discuss the idea that “better outcomes” are reached in “normal deliberative conditions,” neither of these notions is clear. If they reject the very concept of a “good argument,” what kind of standards can be referred to, to decide whether a good final argument has been posed? They just emphasize that deliberating in groups allows people to reach “epistemically superior outcomes and improve their epistemic status.” Mercier (2011a, p. 319) boldly points out: “Here I will have recourse to a more rudimentary yardstick: a good argument is one that is accepted by many people who can understand it and make the effort of evaluating it. Obviously, an argument can be accepted by many people, at a given time, and still be wrong. However, this outcome becomes less likely as the diversity of people who accept it increases.” Because I really do not want to forget my poor critical potential, I could ask: what do you mean by “diversity of people”?

The mere idea of abandoning individuality appears counter-intuitive; we either preserve it or not. M&S point out that “... A distinctive feature of our approach, relevant to the discussion of ‘collective wisdom’, is the claim that the main function of reflective inference is to produce and evaluate arguments occurring in interpersonal communication (rather than to help individual ratiocination)” (Sperber and Mercier, in press). Certainly, in a collective situation all the participants receive the benefits of putting in balance pros and cons of a given argumentation, but did not Robinson Crusoe entertain his brain by imagining worlds until Friday showed up? If we insist that the social condition is the most important for our very nature, which I believe, following Tomasello (2008; Tomasello et al. 2010), then even intuitive inferences are the product of the same social condition.

M&S say, “Reasoning is specifically human. It is clearly linked to language” (Sperber and Mercier, in press). “Clearly linked to language” is not enough, because as far as we know, language is also specifically human, and evolved to make communication more efficient (see Macneilage, 2008). Do reasoning and language then have the same function? Is it because we can represent and verbalize that we can argue? Certainly, my young daughter, who does not talk properly yet, can manifest her disagreement with my orders, but I cannot say that she is having a discussion with me. In the same vein, Mercier (2011a, p. 317) points out that “Cooperation is made more efficient by communication, which in turn is facilitated by the exchange of ar-

guments. Reasoning would have evolved to enable this exchange of arguments.” There is no space here to detail my concern, but I think it is totally the other way round (see Tomasello 2008, p. 105; Tomasello et al., 2010): cooperation is a superstructure that made possible many of our cognitive skills and, certainly, the linguistic recursive faculty as much as the shared human intentionality.

A final epistemological remark is that M&S assume a very classic critical and rationalistic Popperian way of building a theory: “Our definition of reasoning may be debatable, but the argumentative approach to reasoning is about reasoning as we defined it. To object to this definition, it is not enough to offer another definition that may be reasonable and useful. What would have to be shown is that ours fails to identify a phenomenon with enough autonomy and integrity to be a proper object of study and insight” (Mercier and Sperber, 2011b, p. 95), this is to say, they want to be falsified. But, how is one to find a—good—argument to falsify the theory between individuals of a homogeneous group, the group of the scientists?

In this section, I have been trying to advance some observations that arise when the main ideas of M&S's proposal are compared or scrutinised using mainly insights from the field of the psychology of reasoning and to a lesser extent from traditional approaches in argumentation theories.

6. Conclusion

In this review, special concerns have been expressed regarding M&S's reflection on the dual-system theory, epistemic vigilance and confirmation bias, which they use to explain why human beings argue. Because they see reasoning and argumentation as an individual competence which has evolved to serve communicative purposes, from the point of view of argumentation studies—traditionally *dialectics*, *rhetoric* and (*informal*) *logic*—many aspects of their program are still unclear or incomplete. However, M&S's approach should encourage new debates, and for this reason their effort is very much appreciated. Yet it is still necessary to continue improving argumentation studies by adding elements from cognitive science and evolutionary psychology. Perhaps we need to make, or at least try to incorporate a *naturalistic turn* in our field, which is one of Mercier & Sperber's indirect messages.

Acknowledgments: This work is partially a consequence of my post-doc research “Principle of relevance: Cognitive and

argumentative dimensions” (Project No. 3090017, National Commission of Science, Chile, 2008-2010), and part of the project “Flexibility Bias and Perceived Project Success on Software Development: the Chilean Context”, Diego Portales University. I would like to thank the editors of *Informal Logic* who patiently revised this article.

References

- Bernard, S., Mercier, H., and Clément, F. (2012). “The power of well-connected arguments: Early sensitivity to the connective *because*.” *Journal of Experimental Child Psychology*, 111(1): 128-135.
- Berthoz, A. (2009). *Emotion and Reason. The Cognitive Neuroscience of decision making*. New York: Oxford University Press.
- Brown, R. (1999) *Group processes: Dynamics within and between groups*. Oxford: Wiley Blackwell.
- Carey, S. (2009). *The Origin of Concepts*. New York: Oxford University Press.
- Dessalles, J-L. (2007). *Why We Talk. The Evolutionary Origins of Language*. New York: Oxford University Press.
- Dessalles, J-L. (2011). “Reasoning as a lie detection device.” *Behavioral and Brain Sciences*, 34(2): 76-77.
- Eemeren, F. van and Grootendorst, R. (2004). *A Systematic Theory of Argumentation. The Pragma-dialectical Approach*. New York: Cambridge University Press.
- Elqayam, Sh. and Evans, J. (2011). “Subtracting ‘Ought’ from ‘Is’: Descriptivism versus Normativism in the Study of the Human Thinking.” *Behavioral and Brain Sciences* 34(5): 233-248.
- Evans, J. (2009). “How many dual-process theories do we need? One, two, or many?” In J. Evans and K. Frankish (Eds.), *In Two Minds. Dual Processes and Beyond*, pp. 33-54. New York: Oxford University Press.
- Evans, J. (2010). *Thinking Twice. Two Minds in one brain*. New York: Oxford University Press.
- Evans, J. (2011). “Reasoning id for thinking, not just for arguing.” *Behavioral and Brain Sciences*, 34 (2): 77-78.
- Evans, J. and Frankish, K. (2009). *In Two Minds. Dual Processes and Beyond*. New York: Oxford University Press.
- Frankish, K. (2009). “Systems and levels: Dual-system theories and the personal-subpersonal distinction.” In J. Evans and K. Frankish (Eds.), *In Two Minds. Dual Processes and Beyond*, pp. 89-108. New York: Oxford University Press.

- Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press.
- Godfrey-Smith, P. and Yegnashankaran, K. (2011). "Reasoning as deliberative in function but dialogic in structure and origin." *Behavioral and Brain Sciences*, 34(2): 80.
- Govier, T. (1987). *Problems in Argument Analysis and Evaluation*. Dordrecht: Foris.
- Gould, S. (1991) "Exaptation: A Crucial tool for Evolutionary Psychology." *Journal of Social Issues*, 47(3): 43–65.
- Johnson-Laird, P. (2005). "Mental Models and Thought." In K. Holyoak and R. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning*, pp. 185-208. New York: Cambridge University Press.
- Johnson-Laird, P. (2008). *How We Reason*. New York: Oxford University Press.
- Johnson, R. (2000). *Manifest Rationality. A Pragmatic Theory of Argument*. Mahwah, New Jersey: Erlbaum.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Allen Lane.
- Macneilage, P. (2008). *The Origin of Speech*. New York: Oxford University Press.
- Manktelow, K., Over, D. and Elqayam, Sh. (2011). *The Science of Reason. A Festschrift for Jonathan St. B. T. Evans*. Hove: Psychology Press.
- Mercier, H. (2011a). "When experts argue: explaining the best and the worst of reasoning." *Argumentation*, 23(3): 313-327.
- Mercier, H. (2011b). "On the universality of argumentative reasoning." *Journal of Cognition and Culture*, 11(1-2): 85-113.
- Mercier, H. (2011c). "Reasoning serves argumentation in children." *Cognitive Development*, 26(3): 177-191.
- Mercier, H. (2011d). "What good is moral reasoning?" *Mind & Society* 10 (2): 131-148.
- Mercier, H. (in press). "Looking for Arguments." *Argumentation*.
- Mercier, H. and Landemore, H. (Forthcoming). "Reasoning is for arguing: Understanding the successes and failures of deliberation." *Political Psychology*.
- Mercier, H. and Sperber, D. (2011a). "Why do human reason? Arguments for an argumentative theory." *Behavioral and Brain Sciences*, 34(02): 57-74.
- Mercier, H. and Sperber, D. (2011b). "Argumentation: its adaptiveness and efficacy. Response to commentaries on 'Why do human reason?'" *Behavioral and Brain Sciences*, 34(2): 94-111.
- Narvaez, D. (2011). "The world looks small when you only look through a telescope: The need for a broad and developmental

- study of reasoning.” *Behavioral and Brain Sciences*, 34(2): 83-84.
- Oaksford, M and Chater, N. (1998). *Rationality in an Uncertain World*. Hove, UK: Psychology Press.
- Oaksford, M and Chater, N. (2007). *Bayesian rationality: The Probabilistic approach to human reasoning*. Oxford: Oxford University Press.
- Poletiek, F. (2011). “You can’t have your hypothesis and test it: The importance of utilities in theories of reasoning.” *Behavioral and Brain Sciences*, 34(2): 87-88.
- Saunders, C. and Over, D. (2009). “In two minds about rationality?” In J. Evans and K. Frankish (Eds.), *In Two Minds. Dual Processes and Beyond*, pp. 317-334. New York: Oxford University Press.
- Sperber, D. (2001). “An evolutionary perspective on testimony and argumentation.” *Philosophical Topics* 29: 401-413.
- Sperber, D. and et al. (2010). “Epistemic Vigilance.” *Mind and Language*, 25(4): 359-393.
- Sperber, D. and Hirschfeld, L. (2007). “Culture and modularity.” In T. Simpson, P. Carruthers, S. Laurence, and S. Stich (Eds.), *The Innate Mind: Culture and Cognition. Vol. II*, pp. 149-164. New York: Oxford University Press.
- Sperber, D. and Mercier, H. (in press). “Reasoning as a Social Competence.” In H. Landemore and J. Elster (Eds.), *Collective Wisdom*.
- Sperber, D. and Wilson, D. (1995). *Relevance. Communication and Cognition. Second Edition*. Oxford: Blackwell.
- Stanovich, K. (1999). *Who is rational? Studies of Individual differences in reasoning*. Mahwah, NJ: Erlbaum.
- Stanovich, K. (2011). *Rationality and Reflective Mind*. New York: Oxford University Press.
- Stanovich, K. and Toplak, M. (Forthcoming). “Defining features versus Incidental Correlates of Type 1 and Type 2 processing.” *Mind and Society*, DOI, 10.1007/s11299-011-0093-6.
- Stein, E. (1996). *Without good reason: the rationality debate in philosophy and cognitive science*. Oxford: Oxford University Press.
- Stenberg, R. (2011). “When reasoning is persuasive but wrong.” *Behavioral and Brain Sciences*, 34(2): 88-89.
- Sterenly, K. (2003). *Thought in a hostile world: The evolution of human cognition*. Malden, MA: Blackwell Publishing.
- Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: The MIT Press.
- Tomasello, M. et al. (2010). *Why We Cooperate*. Cambridge, MA: The MIT Press.
- Toulmin, S., Rieke, R. and Janik, A. (1979). *An Introduction to Reasoning*. New York: Macmillan.

- Tuomela, R. (2007). *The Philosophy of Sociality. The Shared Point of View*. New York: oxford University Press.
- Walton, D. (2007). *Media Argumentation. Dialectic, Persuasion and Rhetoric*. New York: Cambridge University Press.
- Walton, D. and Krabbe, E.C.W. (1995). *Commitment in Dialogue. Basic Concepts of Interpersonal Reasoning*. New York: State University of New York Press.
- Wiles, J. "Reasoning, robots, and navigation: Dual roles for deductive and abductive reasoning." *Behavioral and Brain Sciences*, 34(2): 92.