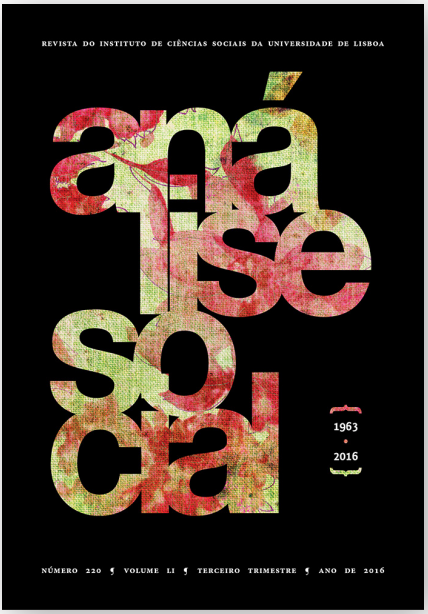

PEDRO JESÚS PÉREZ ZAFRILLA

Is Deliberative Democracy an adaptive political theory? A critical analysis of Mercier's Argumentative Theory of Reasoning

Análise Social, 220, LI (3.º), 2016
ISSN ONLINE 2182-2999

EDIÇÃO E PROPRIEDADE
Instituto de Ciências Sociais da Universidade de Lisboa. Av. Professor Aníbal de Bettencourt, 9
1600-189 Lisboa Portugal — analise.social@ics.ul.pt

Análise Social, 220, LI (3.º), 2016, 544-564

Is Deliberative Democracy an adaptive political theory? A critical analysis of Hugo Mercier's Argumentative Theory of Reasoning. Moral psychology has devalued the role of reasoning in moral life. Nevertheless, the Argumentative Theory of Reasoning (ATR) is a proposal of evolutionary psychology that vindicates the effectiveness of reasoning in communicative processes. This theory even defends its convergence with deliberative democracy. And yet, we hold that the ATR is unsuccessful when supplying a psychological basis for deliberative democracy. The ATR does not adequately explain the epistemic value of deliberation in a way that is congruous with deliberative democracy. To conclude, we will present other shortcomings of the ATR in its concept of reasoning.
KEYWORDS: argumentative theory of reasoning; Hugo Mercier; deliberative democracy; epistemic democracy; reasoning.

É a democracia deliberativa uma teoria política adaptativa? Uma análise crítica da Teoria Argumentativa do Pensamento de Hugo Mercier. A psicologia moral tem desvalorizado o papel do pensamento na vida moral. Não obstante, a Teoria Argumentativa do Pensamento é uma proposta de psicologia evolutiva que reivindica a efetividade do pensamento nos processos comunicativos. Esta teoria defende, inclusive, a sua convergência com a democracia deliberativa. No entanto, neste artigo defendemos que a TAP não é bem sucedida em proporcionar uma base psicológica da democracia deliberativa. A TAP não explica adequadamente a motivação dos sujeitos na deliberação. Para além do mais, a TAP não articula o valor epistémico da deliberação de maneira congruente com a democracia deliberativa. Para concluir, exporemos outras carências da TAP na sua conceção do pensamento.
PALAVRAS-CHAVE: teoria argumentativa do pensamento; Hugo Mercier; democracia deliberativa; democracia epistémica; pensamento.

PEDRO JESÚS PÉREZ ZAFRILLA

Is Deliberative Democracy an adaptive political theory? A critical analysis of Hugo Mercier's Argumentative Theory of Reasoning.

INTRODUCTION[1]

The nature of reasoning has been the subject of constant debate in the realm of moral psychology. In the twentieth century, a rationalistic conception was predominant, thanks to scholars like Piaget, Kohlberg, and Turiel. For them, reasoning constitutes the area of human cognition and the source of moral judgments. However, in the past few decades neurological evidence and the development of evolutionary psychology have given rise to a new perspective. This new model has a broader conception of human cognition. In particular, it includes components like intuition, bias, and heuristics, which lessen the capacity of reasoning in the formation of moral judgements (Le Doux, 1998; Connolly, 2002).

Even so, despite all those differences, both models share the same methodological premise. Reasoning is presented as a cognitive process dedicated to the formation of judgments and reasons through individual reflection. Now, against this approach, a recent psychological theory has risen, labelled Argumentative Theory of Reasoning (herinafter ATR) developed by Hugo Mercier and Dan Sperber among other scholars. It is a stimulating theory of evolutionary psychology that conceives reasoning as a cognitive mechanism developed as part of evolution because of its contribution to the effectiveness of communicative processes.

This emphasis on communication allows Mercier to fix, as one of the aims of ATR the provision of a psychological basis to the ideal of deliberative

1 I acknowledge the support of the Research Project FFI 2013-47136-C2-1-P, funded by the Spanish Ministry of Economy and Competitiveness.

democracy. In particular, ATR attempts to rehabilitate the epistemic value of deliberation from the parameters of evolutionary psychology (Landemore and Mercier, 2012, p. 920). This is no doubt, *a priori*, a complicated task, since neuroethics has sought to question precisely the role of reason in the achievement of good results through deliberation (Haidt, 2012, p.43; Lakoff, 2009, p. 89). And yet, according to Mercier, neuroethics would allow articulating a deliberative model that would lead deliberation to the making of good decisions.

The present paper seeks to analyze ATR from the points of view of moral and political philosophy. We start by expounding the characteristic elements of ATR. Then we delve into the motivations that guide the subjects in the deliberative process, according to this theory. Third, we consider whether or not ATR possesses the correct conception of the epistemic value of deliberation. To conclude, we present the different shortcomings that the ATR model of moral deliberation reveals due to its concept of reasoning. This will allow us to conclude whether the attempt by Mercier to connect evolutionary psychology and deliberative democracy is adequate.

THE ARGUMENTATIVE THEORY OF REASONING

The ATR is characterized by its inconsistency with the concept of the dual process of human cognition, the most common model among scholars of neuroethics (Mercier, 2012, p. 260). For the dual process concept, reasoning constitutes a secondary source of moral judgements, as opposed to intuition. Reasoning represents a conscious, strenuous, and slow mechanism that is only put into effect in case of need. Its function is to modify, alleviate, or where appropriate, replace the influence that intuitions may have upon the formation of moral judgments (Greene, 2007; Haidt 2001; Kahneman, 2012).[2] Consequently, reasoning may form judgements either when it is capable of controlling the impulse of intuitions or when the said impulse has not occurred (Fine, 2006; Pizarro and Bloom, 2003; Saltzstein and Kasachkoff, 2004).

As opposed to this model, Mercier denies that the role of reasoning is the creation of judgements by means of individual reflection. According to Mercier, argumentation is the proper function of reasoning, facilitating interpersonal communication (Mercier and Sperber, 2011, p. 60). The central thesis of the ATR is that the faculty to reason was developed in the course of evolution

2 As to the role allotted to both cognitive processes, scholars disagree. Greene holds that intuition lies in the genesis of our deontological judgements and the cosequentialist judgements would have an rational origin. On the contrary, Haidt is of the opinion that the majority of our judgements are intuitive.

thanks to its adaptive advantages within the framework of inter-subjective relationships. The exchange of information in the argumentative process will provide manifold advantages such as the maintenance of personal reputation, the solution of any type of problems, the detection of *free riders* by means of gossip, or the making of better decisions, for example. Thus, the role of reasoning in the communicative processes will be that of contributing to the effectiveness of the said communication by discriminating between trustworthy and untrustworthy information.

According to the ATR, reasoning is composed of two cognitive processes. Mercier defines reasoning as "a specific cognitive mechanism that aims at finding and evaluating reasons, so that individuals can convince other people and evaluate their arguments" (Landemore and Mercier, 2012, p. 914; see also Mercier and Sperber, 2011, p. 59). Such processes are well differentiated and fulfil different functions. In the first place that of paying attention to and evaluating the reasons for and against the matter within the framework of the deliberation process. To be more precise, in evaluation, the person attempts to analyze the arguments of the interlocutor in order to assess the trustworthiness of the person and of the information provided by him. In consequence, he could decide whether to accept its contents or not.[3] In this way, the arguments of the interlocutor allow the person to recognize errors in his perception of reality and to reconsider his own position (Mercier, 2011, p. 143; Landemore and Mercier, 2012, p. 915). On the other hand, and versus this evaluative function of reasoning, there is another function: to find and articulate arguments,[4] with a double objective. The first is to respond to the arguments of the opponent when they are not considered conclusive. The second is to attempt to convince the interlocutor of one's stance. This double evaluation and argumentation process will let the person make better decisions within the framework of communicative processes (Mercier, 2011, p. 132).

Now, this characterization of reasoning brings the ATR closer to the theory of the dual process of human cognition. On the one hand, the ATR coincides with Greene in its vindication of the faculty of reasoning in inter-subjective processes. In this particular aspect, the ATR opposes Haidt's postulate that deliberation is no more than a process of emotional persuasion (Paxton and Greene, 2010; Greene, 2013, p. 335; Haidt, 2001). But, in contrast to Greene, the ATR shares with Haidt the idea that reflection, with exceptions, is only fruitful

3 This information discrimination is carried out by means of the cognitive mechanism of epistemic vigilance developed in the evolution. According to the ATR scholars, the aim of epistemic vigilance is to allow the survival of individuals in communicative contexts (Sperber, 2010).

4 This function is hereinafter labeled "arguing".

when it is carried out according to the information provided by third parties. Thus, the ATR defends that reasoning always has a social and argumentative function. The position of the ATR in relation to the theories of the dual process of reasoning can thus be expressed with a Mercier citation. For him, ATR "can be seen as a refinement of Haidt's Social Intuitionist Model" (Mercier, 2011, p. 132) since this theory grants more prominence to reasoning than Haidt's, even though it retains the social character of reasoning.

This said, reasoning will adequately fulfil the function of ameliorating the epistemic effectiveness of communicative processes if developed within adequate conditions, or, as Mercier calls them "normal conditions". To be precise, there are basically three conditions. The first is that reasoning be developed in the framework of the exchange of conflicting opinions, as far as the subject matter under debate is concerned (Mercier and Landemore, 2012, p. 248; Thompson, 2008). The second is that an evaluation of the interlocutor's arguments takes place (Mercier, 2011, p. 140; Landemore and Mercier, 2012, p. 916). Third, the participants should be adequately motivated (Merier and Sperber, 2011, p. 61). That is, they "should be truth-oriented" and should not attempt to defend a recalcitrant stance in the dialogue. In these conditions, the exchange of arguments will result in the making of better decisions by means of information evaluation. In other words, these conditions give shape to a process of deliberation as the ATR understands it.

On the other hand, according to the ATR, individual reflection has little epistemic efficiency (Mercier and Sperber, 2011). The reason for this is that personal reflection is largely a kind of biased reasoning (Lord, Ross and Lepper, 1979; Kunda, 1990). Usually, in personal reflection, reasoning does not neutrally evaluate the evidence present in the subject matter. Indeed, reasoning is reduced to a mechanism of argument creation, in accordance with the previous attitudes and impressions that the person has regarding the subject matter at hand. This is why the person will be able to approach the subject matter in a non-biased form, thanks to the information presented by an interlocutor he disagrees with (Landemore and Mercier, 2012). Hence, for the ATR and against the tenets of the dual process theory, the proper function of reasoning is the evaluation of information supplied by others, not that of reaching an opinion in private.

Likewise, according to Mercier, the exchange of opinions between like-minded people and the dialogue between persons who recalcitrantly continue to defend a stance are non-deliberative forms of reasoning. This is because the conditions indicated above are not met: either there is no diversity of opinions or the subjects are not truth-oriented (Mercier and Landemore, 2012). Above all, there is no evaluation, but mere production of arguments. Thus, these forms of communication should be better labelled conversations or discussions, but

not deliberation (Mercier and Landemore, 2012). In them, the dialogue context unleashes a set of cognitive biases, such as confirmation, as well as a tendency to desire to be perceived favorably by the group or the reaffirmation of non-deliberative opinions (Schkade, Sunstein and Hastie, 2010). These biases will drive reasoning to a polarization of stances instead of to an agreement.

This is a thought-provoking proposal, indeed, and yet it is not lacking in problems. In the following pages, we shall expound upon those aspects that we consider most arguable in the ATR.

THE MOTIVATION OF SUBJECTS IN THE DELIBERATIVE PROCESS

The first difficulty lies in the motivation that guides persons in the deliberative process, according to the ATR. In principle, according to Mercier's approach, the motivation of people in arguing is to try to convince the interlocutor of the theses defended. And yet, this motivation should not be understood as an attempt to convince the other person of what one considers to be just (Landemore and Mercier, 2012). This is so because the ATR, as an evolutionary-psychology proposal, results in the argument that reasoning evolved due to its usefulness to facilitate the adaptation of individuals to inter-subjective contexts. In consequence, we can deduce that the aim of the arguer is to project a good image of himself before the others, since that will facilitate his survival within the group. In this way, and in consonance with other scholars, Mercier ascribes to the subjects within a deliberation the motivation of upholding one's reputation inside the group (Trivers, 1971; Alexander, 1987; DeScioli and Kurzban, 2009; Sperber and Baumard, 2012). We can observe this most especially in Mercier's analysis of deliberation dealing with moral and political subjects. In deliberation, the individual will offer adequate arguments not to defend what he considers just, but in such a way that the others might have a favorable image of him and, by extension, of what he thinks about the subject matter (Mercier, 2011). This is because a negative image might derive from rejection by the others.

This search for reputation as the goal of arguing is shown in the "objective criteria" that Mercier (2011) fixes for moral valuation. The said criterion consists of the clear cut image that the group has of an individual in particular. But this thesis of Mercier's is, without any doubt, problematic and doubly so, first because for Mercier, moral "objectivity" is limited to consensus in the opinion that the group might have of a particular individual that determines his reputation[5] (but, as we shall show below, group loyalty and moral values are

5 Hence, the motivation of the subject to construct a good image of himself in the course of argumentation.

two different realities). Second, Mercier uses the concept of moral objectivity in a way which is inexact. Today, the mere fact of defending the existence of objective moral evaluation or of "objectivity" within morals is problematic in itself, due to the prevailing moral pluralism. Instead of objectivity —and moral relativism— it seems much more adequate to focus on moral evaluation from a presumption of universality as an element inherent to moral judgements (Habermas, 1997).

This orientation of arguing to uphold group survival is also present in another function that Mercier (2011) grants to reasoning. It consists of warning the other person about the negative consequences that a behavior that is contrary to the principles he purports to defend would hold in store for his well-being. In this case, the interlocutor will acknowledge the superiority of the said arguments since —within the evaluation process— his aim is to uphold his reputation within the group. Indeed, Mercier points out that the person will allow himself to be convinced as a result of the pain perceived when recognizing the consequences it might have for him, should society discover his hypocrisy. This sensation of pain makes a person change opinion in the course of a dialogue because of the rejection generated in his mind when faced with the possibility of being considered inconsistent.

Thus we see that for Mercier, as for other neuropsychology scholars, the foundation of moral conscience lies in an adaptive code formed in our brain throughout evolution as a result of living together in small groups. This code triggers intuitive responses of either pleasure or pain, depending on the influence that the specific behavior might have on well-being. Thus, this evolutionary code produces behavior that is appropriate in order to promote survival within the group (Haidt and Bjorklund, 2007; Lakoff, 2009; Westen, 2007). Moral consciousness would then be a source of prudential motivation, not moral motivation, nor epistemic process, e. g. one that aims at truth. The neuro-biological reality of that motivational force would consist of the sensation of the pain that appears when we recognize the consequences that our behavior would produce for our reputation within the group. The only difference between Mercier and other scholars like Lakoff, Haidt, and Western is the fact that Mercier assigns to reasoning an efficiency value that is denied by these other scholars and which they assign to the metaphors in its stead. Obviously, a central subject here would be the proof of whether or not we can really reduce moral conscience to that strategic voice that helps us sustain reputation (Cortina, 2013), but, this goes beyond the limits of the present paper, centered as it is on deliberation.

If according to the ATR the real motivation for the subjects in deliberation is to argue properly in order to generate a good image *vis-à-vis* the others

and so maintain reputation, then deliberation will consist of a procedure by which the subjects will attempt to make prudential decisions. So, for example, when the subjects take part in a debate on moral or political issues, their arguments could be directed to defend a stance that will preserve their good image within the group. On the contrary, to defend a stance that is critical of what is socially established would imply negative consequences for their well-being. In consequence and according to Mercier, the argumentative process would always be secondary to the strategic calculation of what can be said and what cannot be said to safeguard reputation and avoid unpleasant situations.

This seems to prove that Mercier does not have an adequate concept of moral deliberation. This is firstly because it is difficult to reduce moral deliberation to the presentation of prudential reasons. According to the Kantian approach, it is easy to recognize that the said arguments directed to the maintenance of reputation are not actual moral arguments but hypothetical pieces of advice that appeal to the strategic voice that Alexander speaks of (Kant, 1785/1964). In consequence, anyone who argues with the intention of maintaining his reputation (or with the intention of helping another maintain his) does not have a real moral but a prudential intention and he does not take part in doing so in a moral but in a strategic rationality (Habermas, 1995). This is because he does not attempt to defend what he considers to be just, but to guarantee for himself -or the other- survival within the group. As proof thereof, in order to attain that goal the recourse to lying or to rhetoric could be deemed legitimate. And yet, as we shall see below, these recourses are banned within the framework of moral deliberation from the point of view of deliberative democracy.

Also, to reduce motivation to the search for reputation leads the ATR to deny people the critical capacity to recognize the immorality of social conventions. Truth orientation would, thus, be reduced to the respect for social convention to ensure survival. And yet this does not correspond to the common concept of deliberation, especially in deliberative democracy.[6]

THE EPISTEMIC VALUE OF DELIBERATION

The second problem specifically affects the defense by Mercier of the epistemic value of deliberation within the ATR and the convergence of this theory with deliberative democracy. Mercier is categorical:

6 This matter will be approached below.

> the argumentative theory can make predictions regarding deliberation bearing on any kind of questions that can be assumed to have a better or worse answer; whether from a factual, moral or political point of view [Mercier and Landemore, 2012, p. 245].

> even if the function of reasoning is argumentative –to produce and evaluate arguments- rather than purely epistemic, it should still lead to an improvement in epistemic status [Landemore and Mercier, 2012, p. 919].

> the argumentative theory of reasoning is thus not only compatible with an epistemic deliberative approach but, as this paper argues, its predictions converge with the predictions of deliberative democrats [Landemore and Mercier, 2012, p. 920].

To prove the convergence of the ATR with deliberative democracy, Mercier assimilates information evaluation within the dialogue –carried out by the cognitive process of evaluation– with the acknowledgement of the better argument as held by Habermas (Landemore and Mercier, 2012, p.920). Nevertheless, the ATR theses are not convergent with deliberative democracy. The reason is because this adaptive theory has too narrow a concept of the epistemic value, which does not correspond to the one developed by the parameters of deliberative democracy.

Mercier interprets the epistemic value of deliberation in an evolutionary sense. The deliberative processes among individuals oriented by the prudential norm of cooperation give better results in the sense that they will allow the upholding of reputation and, with it, the survival of the persons within the group. Properly understood, the moot point is whether the epistemic value of deliberation can be reduced to the result of a sensible norm dictated by evolution, in other words, if the individuals judge a decision as good or bad depending on whether the result is convenient or not as a means to uphold their survival**.** To analyze this point, it will be convenient to review what has been indicated by the scholars of epistemic value in deliberative democracy.

Deliberative democracy is a normative theory that emerged in the 1980s in opposition to the existing liberal democracy, which had been plunged into a crisis of confidence and citizen apathy. Deliberative democracy developed out of different theoretical approaches: republicanism (Barber, 1984), liberalism (Rawls, 1993; Gutmann and Thompson, 1996; Richardson, 2002; Crocker, 2008) and discourse ethics (Habermas, 1997; Bohman, 1996; Benhabib, 1996; Cortina, 1985). Other authors have also developed various proposals for deliberative democracy beyond these approaches, such as Nino (1996), Fishkin (1995), and Fung (2005). In spite of these differences, all the proposals share some characteristic features opposed to the liberal model: in opposition to the

aggregation of fixed preferences, deliberative democracy will take a position in favor of the transformation of preferences; in opposition to the satisfaction of private interests, a joint search for the common good is encouraged; and in opposition to the liberal mere secret vote, accountability and public justification are demanded. But above all, in opposition to a liberal system based on negotiation between interests, deliberative democracy makes public deliberation the axis of the political system. Deliberation is generally understood as discussion based on rational argument, the exchange of information and the triumph of the best argument. The goal of deliberation is to reach an agreement on the best decision, independent of the private interests of the individuals. Nevertheless, the heterogeneous origin of this political theory has led to its opposition to liberal democracy being moderated with the passage of time (Bohman, 1996). In this way, the initial confrontation between deliberation and negotiation has given way to recognition of difference, of emotions, and of other forms of communication, such as storytelling (Chambers, 2003; Bohman, 1996).

The defense of the epistemic value of deliberation is another shared thesis among scholars of deliberative democracy and constitutes one of the central pillars of this political theory (Bohman, 1996; Freeman, 2000; Pérez Zafrilla, 2009). In this democracy theory the element that directs the deliberative process toward an optimal decision is not sensible calculation, oriented toward survival within a socially established context. On the contrary, within democracy the epistemic value is defined in terms of correction or impartiality (Cohen, 1986; Martí, 2006). According to this political theory, when citizens take part in deliberation they do so guided by a criterion of justice. That is, those who defend a dialoguing stance do so because they believe it is the best one or, at least the better of two alternatives, not because they are satisfying their own interests. This criterion of justice has two fundamental elements: in the first place, it is inter-subjective, since the subjects attempt to make the others recognize that their position is the right one; on the other hand, it is independent of personal interests (Elstub, 2010). Each subject expounds his arguments with the intention of rationally convincing the others, and at the same time accepts being convinced by a better argument that allows everyone to see, by means of reasons, that that proposal is the better one. In this way, and within this epistemic concept, the deliberative process is characterized by being a *belief adjustment* about what constitutes the best option to take by recognizing the best argument.

Consequently and in accordance with this epistemic concept, the motivation that drives deliberation is not the maintenance of reputation with the group, as Mercier suggests, but the defense of what is considered just. Now that

motivation includes a key element not identified by Mercier: an ethical attitude, a disposition to recognize a moral relationship among subjects who must be respected (Cohen, 1997). Indeed, the establishing of a deliberative process about what is just with another person implies the attitude of recognizing the other person as a valid interlocutor. The communicative link forces the subjects to listen and to take into account the needs of the others, to reflect on how those needs could be met and, finally, to reach a common solution by means of dialogue. However, that dialogue also presupposes recognition among the participants within a symmetrical relationship (Habermas, 1990). This implies that all its members can participate on an equal footing and that the process is founded upon rational arguing. Each subject propounds his arguments under the pretense of rationally convincing the others of the validity of his proposals, and at the same time accepts being convinced by a better argument that will allow him to see that the other proposal is the correct one. At the end, this will only be possible if, within the deliberation, all participants acknowledge themselves to be guided by a criterion of what is just (Habermas, 1997).[7]

Thus, these epistemic and ethical elements are the conditions that embody the deliberative processes and not the mere formal conditions pointed out by Mercier —the disagreement among participants or an attitude directed to the maintenance of survival. This is so because deliberation is not directed to the attainment of agreements that will allow the maintenance of survival within a group of specific individuals. Deliberation aims to establish an agreement about what is considered just on a contrafactual level, beyond the socially established frameworks and the private interests of the subjects. This epistemic concept of arguing can also account for morality. Indeed, moral order constitutes a critical sphere that allows us to transcend what is socially established and to rationally evaluate whether or not the social convention matches principles deemed universal (Cortina, 2010). On the contrary, the concept propounded by Mercier consists of hypothetical pieces of advice directed to the attainment of social approval. As a consequence, the alleged convergence between the ATR —to produce and to evaluate arguments— and deliberative democracy —epistemic in character— does not take place. This is so because

7 Other authors have accentuated the presence of other types of motivations in public debate, such as the expression of personal experiences for women, the expression of religious convictions for believers, and the defense of particular interests of minorities such as homosexuals, immigrants, or people with disabilities (Sanders, 1997; Chambers, 2003). Nevertheless, all of these motivations still retain an ethical component: the ethical demand for equal recognition by the other people in the public forum (Cortina, 2007). Therefore, from different perspectives and with different voices, their objective remains that the other citizens recognize their demands as just.

Mercier bases his deliberation model upon a strategic rationality that is guided by the survival interest, not upon a properly moral rationality that would favor rational agreements on what is just. The latter is what deliberative democracy sustains (Benhabib, 1994; Markovits, 2006).

REASONING

One last problem with the ATR is that it does not carry out an adequate analysis of moral deliberation. According to Mercier the two cognitive mechanisms that give form to reasoning are dissociated.

> there is an asymmetry between the *production* of arguments, which involves an intrinsic bias in favour of the opinions or decisions of the arguer whether they are sound or not, and the *evaluation* of arguments, which aims at distinguishing good arguments from bad ones [Mercier and Sperber, 2011, p. 72].

We can deduce from this that only the evaluation mechanism *in the interpersonal context* between opposed stances will maintain its rational character. This evaluation will also allow the epistemic amelioration of the deliberative processes in the sense pointed out by Mercier. In the meantime, the process of arguing will not be directed to the presentation of reasons after a period of reflection, nor will it aim to reach a reasonable agreement that will bring the stances of what each party considers to be just closer. Quite the contrary, arguing merely seeks and presents reasons that match the stance of the arguer, motivated by the opinion the person holds on the subject matter. That is, arguing would consist of the creation of *post hoc* arguments of a biased nature, for example, Haidt's Link 3 of Social Intuitionist Model (Haidt and Bjorklund, 2007). In fact, Mercier and Sperber (2011) even deny that the arguments are the result of System 2, but say that they would derive from the intuitive inference belonging to System 1.

This turns reasoning into a partly reflexive —evaluation— and partly intuitive —arguing— process. Its main consequence will be that in deliberation an indirect agreement is reached related to the best decision, by means of a balancing of information biases of confirmation that are present in the arguing of the interlocutors. That balance is carried out by reciprocal evaluation and the counterarguing that each subject carries out with the biased arguments of his interlocutor. The process allows each one to change his position until a shared opinion is achieved. In this sense, Mercier indicates that in the deliberation the truth-oriented subject should counteract the confirmation bias present in the interlocutor's argument (Landemore and Mercier, 2012).

Still, it seems evident that this presentation of the deliberative process has little in common with real life. To be more specific, this approach on reasoning poses at least two problems. The first is that a dissonance between reasoning in the processes of evaluation and arguing is produced. The second is that Mercier does not perceive the complexity of arguing.

DISSONANCE OF REASONING

Mercier presents the processes of arguing and of evaluation as two disconnected psychological mechanisms, resulting in serious dissonance between them. True, Mercier insists that his approach is merely descriptive and centered only upon the psychological aspects of deliberation (Mercier and Landemore, 2012). However, this dissonance between arguing and evaluation results in unacceptable normative implications.

One can accept, in principle, that reasoning may fulfil different functions —evaluation and arguing— and that both functions might have different underlying cognitive mechanisms. But that cannot lead us to conclude that there is no connection between both cognitive processes in such a way that one might respond to a truth orientation while the other one would be biased. Affirming that there is dissonance between the processes of evaluation and persuasion introduces a sort of schizophrenia within the deliberative process. Indeed, according to Mercier's approach the subjects would act in the following way: they possess a truth-orientation that enables them to critically evaluate whether their interlocutor is right, they conclude that he is wrong and decide to counterargue. Nonetheless, what happens then is that the reasonings elaborated by the subjects are the product of a confirmation bias and, hence, alien to this truth-orientation.

But this approach by Mercier is erroneous. Deliberation implies a much more complex process than the one proposed by the ATR. In fact, reasoning cannot be adequately understood with mere psychological processes in mind (Monin, Pizarro, and Beer, 2007). Reasoning is a complex process in which it is not possible to split evaluation and arguing as the ATR does. The reason is that in deliberation the individuals evaluate the arguments of the interlocutor with a set of beliefs and values that are assumed to be valid. Those assumed beliefs and values make up the cognitive framework that articulates the concept that each individual has of reality (MacIntyre, 1981; Gutmann and Thompson, 1996; Gaus, 1997). As the communitarian scholars defend, that reference framework of beliefs shapes the very identity of the subjects (Taylor, 1989). In this sense, and as Alston (1985) points out with his idea of epistemic justification, a person is epistemically justified to believe something as long as she bases her evaluation on the pieces of evidence that she considers to be just.

A consequence thereof is that a person cannot keep distance from her assumed beliefs to check whether her concept of reality is adequate or not. This is why, contrary to Mercier's claims, the arguing process cannot be considered biased if the subjects do not stop to consider whether their propositions are correct or not (Landemore and Mercier, 2012). Carrying out such verification would require placing oneself at a vantage point that is independent of the self-same beliefs that make up the evaluation of the subject. The reason for this is that both the evaluation and the arguing participate in the same belief grounds that articulate the "framework of reference" of the person. As a matter of fact, the reasonings produced by the arguing subjects, far from being *post hoc* reasonings, derive from a reflection that stems from the set of beliefs adopted in society (Levi, 2007). In other words, when someone concludes that his interlocutor is right, the reasonings that he articulates in order to try to convince him would be those considered adequate according to the same criterion used to reject the reasonings of his interlocutor, since, as presented in the previous section, the subjects attempt to persuade the others of what they consider just (or true).

Likewise, the fact that the subjects will choose the information favorable to their thesis while rejecting the contrary does not constitute a confirmation bias. Actually it is a perfectly natural process. In fact, it is surprising that Mercier should defend the existence of the confirmation bias, when one can deduce, from the epistemic vigilance thesis developed by Sperber and Mercier, a rejection of the relatedness bias. According to the concept of epistemic vigilance, a person is epistemically justified to ascribe credibility to others according to the degree of trust that she has in the tenet that the opinion of the others can guarantee her survival (Sperber, Clément, Heintz, Mascaro, Mercier, and Origgi, 2010). That is, people tend to rationally trust the people about whom they have proof sufficient to conclude that they favor their survival (Liao, 2011; Jacobson, 2012).

But if people are epistemically justified to grant credibility to those they trust, the same can be said of the confirmation bias. To grant a higher degree of credibility to such pieces of evidence as coincide with our own and to mistrust those that contradict them or, similarly, to center upon refuting what we do not believe instead of questioning what we do believe does not correspond to a cognitive bias either. On the contrary, it consists of a fully rational attitude, since all our approximations to reality stem from a comprehension framework built upon specific pieces of evidence. What would be irrational would be for people to doubt the pieces of evidence that coincide with their *Weltanschauung* (Liao, 2011). When somebody is epistemically justified in believing something —even if that something is wrong— it does not make

sense to attempt to demand an attitude of skepticism toward his concept of reality, since to demand this from him would imply, as Mercier says, asking him to place himself at a vantage point alien to his individual comprehension framework. That is, it would be like asking the person to hold a schizophrenic attitude toward reality.[8]

Going back to arguing, the fact that it might consist of the articulation of arguments that are favorable to the thesis defended by the subject also does not make the process biased. This way of proceeding is fully rational, insofar as the selection of information to articulate arguments is based upon what the people consider correct according to their system of beliefs. In arguing, people articulate favorable arguments because those are precisely the ones they believe in and they reject the arguments that do not coincide with their beliefs because they do not consider them adequate. The creation of arguments in the deliberative context is not biased, because the person cannot leave her cognitive framework to check whether the reasons she is articulating are adequate or not to convince her interlocutor. She could only know that if she had access to the cognitive framework of the interlocutor, but that is impossible. The subject can articulate reasons stemming only from the information and beliefs of her own cognitive framework. Thus, if someone believes p, it is only normal that she will articulate reasons that are favorable to p. The only properly irrational behavior would be to believe p after a truth-oriented process and bearing in mind the reasons posed by her interlocutor and elaborating reasons against p. This is so because the arguments articulated by people are based upon what they consider to be just.

In consequence, to defend asymmetry between evaluation and arguing on the basis that both are based upon different cognitive mechanisms is not acceptable. Because we cannot conclude from the psychological bases —psychological processes— that evaluation and arguing do not stem from the same set of beliefs and that they share truth-orientation. Moreover, as we pointed out in the previous section, arguing has the same degree of epistemic value as evaluation and it is not possible to split the two processes. In consequence, far from being a biased process, arguing is directed to the defense of theses that are founded upon the beliefs that articulate the individual's concept of reality. Finally, this idea connects with the arguing model developed by Mercier.

8 A different matter is the reasonable attitude that is necessary for the person to recognize the fallibility of her beliefs and the readiness to let oneself be convinced by arguments that prove to be better during the process of dialoguing, since, without that reasonable attitude deliberation would simply be unfeasible (Cohen, 1997).

COMPLEXITY OF ARGUING

Mercier does not understand arguing adequately. In the first place, the dissonance between evaluation and arguing that he postulates prevents him from recognizing the point at which his proposal and deliberative democracy might converge. Mercier presents arguing as a biased process and rejects the idea that the reasoning subjects might attempt to present "the better argument." In his point of view, arguing is governed by the heuristics of minimal effort, according to which people halt their reasoning at "the first argument that comes to mind" and will look for another only if this one is refuted by their interlocutor (Mercier, 2011). This approach is short-sighted precisely because the search for the better argument and the heuristics of the minimal effort are two perfectly compatible theses. One of them is epistemic while the other is psychological. Indeed, when person A argues, he attempts to present the best arguments possible to rationally convince his interlocutor B. To do so, he will argue from a criterion of what is just until he arrives at the first argument he holds appropriate. On the other hand, interlocutor B will, in turn, evaluate the arguments of A and try to convince him according to what B considers just. Then, A will try to articulate new, different arguments he considers valid to refute the arguments set forth by B. This process will be repeated until an agreement is reached that puts an end to the deliberative process. Consequently, the fact that deliberation might consist of an exchange of arguments based upon the heuristics of minimal effort does not imply that the process of reasoning is not based upon the criterion of what is just nor that the persons might not be willing to let themselves be persuaded by recognizing the better argument. The defense of epistemic deliberation by deliberative democracy serves to reinforce this contradistinction (Cohen, 1997; Gutmann and Thompson, 1996).

On the other hand, arguing is a fully evaluative process. In the elaboration of reasons, people take epistemic and moral elements into account to find reasons that are not only effective but also acceptable. So, for instance, the recourse to demagoguery would be highly effective to make the interlocutor have a good image of the speaker. And yet, demagoguery or fallacies are resources that are excluded from deliberation. The reason is that they violate the moral duty of recognizing the interlocutor as a subject worthy of respect who should be convinced by means of reasons she can reasonably accept (Richardson, 2002; Rawls, 1993). Thus, in arguing, the interlocutors start off from a presupposition that the recourse to demagoguery is morally unacceptable. In consequence and through a process of reflection they will exclude the set of reasons based upon fallacies or sentimental demagogical appeals (Cohen, 1997, Benhabib, 1996). This shows that arguing is a controlled psychological process, like the formation of judgements itself (Fine, 2006).

Finally, this conjunction of epistemic and moral parameters also delimits the arguments that are deemed acceptable within the different deliberation contexts. This is because different contexts require the use of different types of reasons (Baier, 1969). So, for example, the public justification of norms requires that the reasons set forth fulfill two prerequisites. On the one hand, they should be reasons that every citizen would reasonably accept (Gaus, 1991; Richardson, 2002). That is, they should be reasons based upon shared moral values (Gaus, 1997). On the other hand, and as a consequence, those reasons should be sincere. That is, those reasons should really motivate the subject to defend the theses he defends. Otherwise it would be hypocrisy, something strongly criticised by both the Liberals and the scholars of deliberative democracy (Rawls, 1993; Gutmann and Thompson, 1996).

On the other hand, when one argues in a different context the demands change. Thus, when arguing is carried out with the intent of convincing the interlocutor to support a proposal, the arguer should bear in mind what is persuasive for his interlocutor, not what actually motivated him to support the measure. In order to do so, he will have to articulate reasons that may be convincing for the interlocutor, even though they might not be convincing for the arguer himself (Weithman, 2002). For example, an atheistic politician could employ religious arguments to convince faithful Presbyterians to support her proposal. In this case, the reflection on the appropriate reasons will disregard the principle of sincerity, since the goal is to appeal to the framework of reference of the interlocutor to make him change his mind on the subject. Therefore, it will not be immoral, in this case, to make use of reasons one does not believe in (Gaus, 1997).

In this way, to use Mercier's terms, calm and evaluative reflection and epistemic value are present not only in the scrutiny of the interlocutor's arguments but also in the elaboration of one's own arguments. This is because the aim of arguing is to articulate the appropriate arguments to achieve the proposed objective within each situation (public justification or persuasion), carried out from a specific cognitive framework. In consequence thereof, deliberation represents a complex process that cannot be dissected, as the ATR attempts, like a series of cogs in a set of psychological mechanisms.

CONCLUSION

The ATR represents a peculiar theory of moral psychology that sets itself apart from the proposals presented by the dual concept of human cognition. Patently an heir to social intuitionism, its main advantage with regard to other proposals is its acknowledgement of the role of reasoning within the deliberative

process. Even so, this model has to confront certain obstacles. In the present paper we have presented a series of critiques, referring, in particular, to the purposes of this psychological theory of being a descriptive proposal not only compatible with but even convergent with deliberative democracy. All of these critiques have the denouncement of the reductionist focus that characterizes the ATR as a common denominator. It is difficult to vindicate the role of reasoning in deliberation when this theory is incapable of recognizing reasoning in its complexity.

Within ATR, reasoning is conceived as a schizophrenic mechanism composed of two psychological processes in which one of them, evaluation, has the role of correcting the biased procedure of the other, arguing. This, coupled with the adaptive function granted to reasoning, is precisely what impedes this theory from recognizing the epistemic value of deliberation. Furthermore, we can infer from Mercier's argument that the goal pursued by the subjects in the processes of moral deliberation is to guarantee their reputation before others. But this interpretation is far removed from the motivation that characterizes the deliberative processes for the scholars of deliberative democracy.

We thus observe that the reductionist approach of the ATR makes this theory incapable of recognizing the complexity that characterizes reasoning and deliberation. This is because reasoning is no mere succession of psychological processes, nor is deliberation characterized by the development of reasoning in conditions under which disagreement is a matter of formality. On the contrary, reasoning is a process that is perfectly integrated in evaluation and arguing. Moreover, deliberation has epistemic and moral dimensions that are lacking in Mercier's approach. To sum up, it seems reasonable to conclude that the ATR not only has an erroneous concept of deliberation, but also fails in its attempt to propound a psychological basis for deliberative democracy.

REFERENCES

ALEXANDER, R. (1987), *The Biology of Moral Systems*, New York, Aldine de Gruyter.

ALSTON, W. P. (1985), "Concepts of epistemic justification". *Monist*, 68 (1), pp. 57-89.

BAIER, K. (1969), *The Moral Point of View: a Rational Basis of Ethics*, Ithaca and London, Cornell University Press.

BARBER, B. (1984). *Strong Democracy: Participatory Politics for a New Age*, London, University of California Press.

BENHABIB, S. (1994), "Deliberative rationality and models of democratic legitimacy". *Constellations*, 1 (1), pp. 26-52.

BENHABIB, S. (1996), "Toward a deliberative model of democratic legitimacy". *Democracy and Difference*, Princeton, Princeton University Press, p.83.

BOHMAN, J. (1996), *Public Deliberation*, Cambridge, The MIT Press.

CHAMBERS, S. (2003). "Deliberative Democratic Theory". *Annual Review of Political Science*, 6 (2), pp.307-326.

CONNOLLY, W. E. (2002), *Neuropolitics. Thinking, Culture, Speed*, Minneapolis, University of Minnesota.

COHEN, J. (1986), "An epistemic conception of democracy". *Ethics*, 97 (1), pp. 26-38.

COHEN, J. (1997), "Deliberation and democratic legitimacy". *In* J. Bohman, W. Rehg (eds.), *Deliberative Democracy. Essays on Reason and Politics*, Cambridge, The MIT Press, pp. 67-92.

CORTINA, A. (1985), *Crítica y Utopía: la Escuela de Francfort*, Madrid, Cincel.

CORTINA, A. (2007), Ética de *la razón cordial. Educar en la ciudadanía en el siglo XXI*, Oviedo, Nobel.

CORTINA, A. (2011), *Neuroética y Neuropolítica. Sugerencias para la educación moral*, Madrid, Tecnos.

CORTINA, A. (2013), "La conciencia moral: Entre la naturaleza y la autonomía". *Cuadernos salmantinos de filosofía*, 40, pp. 249-262.

CROCKER, D. A. (2008), *Ethics of Global Development: Agency, Development and Deliberative Democracy*, Cambridge, Cambridge University Press.

DESCIOLI, P., KURZBAN, R. (2009), "Mysteries of morality". *Cognition*, 112 (2), pp. 281-299.

ELSTUB, S. (2010), "The third generation of deliberative democracy". *Political Studies Review*, 18 (3), pp. 295-296.

FINE, C. (2006), "Is the emotional dog wagging its tail or chasing it?". *Philosophical Explorations*, 9, pp. 83-98.

FISHKIN, J. (1995), *The Voice of People*, London, Yale University Press.

FREEMAN, S. (2000), "Deliberative democracy. Sympathetic comment". *Philosophy and Public Affairs*, 29 (4), pp. 371-418.

FUNG, A. (2005). "Deliberation before the Revolution. Toward an Ethics of Deliberative Democracy". *Political Theory*, 33 (2), pp. 379-419.

GAUS, G. (1991), "Public justification and democratic adjudication". *Constitutional Political Economy*, 2 (3), pp. 251-281.

GAUS, G. (1997), "Reason, justification and consensus. Why deliberative democracy can't have it all". *In* J. Bohman, W. Rehg (eds.), *Deliberative Democracy. Essays on Reason and Politics*, Cambridge, The MIT Press, pp. 205-242.

GOODIN, R. (2005), *Reflective Democracy*, Oxford, Oxford University Press.

GREENE, J. (2007), "The secret joke of Kant's soul". *In* W. Sinnot Armstrong (ed.) *Moral Psychology*, vol. 3, Massachusetts, The MIT Press, pp.35-79.

GREENE, J. (2013), *Moral Tribes. Emotion, Reason, and the Gap between us and them*, London, Atlantic Books.

GUTMANN, A., THOMPSON, D. (1996). *Democracy and Disagreement*, Cambridge, Belknap Press of Harvard University Press.

HABERMAS, J. (1990), *Moral Consciousness and Communicative Action*, Cambridge, Polity Press.

HABERMAS, J. (1995), *Justification and Application. Remarks on Discourse Ethics,* Cambridge, Polity.

HABERMAS, J. (1997), *The Theory of Communicative Action*, Cambridge, Polity Press.

HAIDT, J. (2001), "The emotional dog and its rational tail. A social intuitionist approach to moral judgement". *Psychological Review*, 108 (4), pp. 814-834.

HAIDT, J. (2012), *The Righteous Mind. Why Good People are Divided by Politics and Religion*, New York, Pantheon Books.

HAIDT, J., JOSEPH, C. (2004), "Intuitive ethics: How innately prepared intuitions generate culturally variable virtues". *Daedalus*, 133 (4), pp. 55-66.

HAIDT, J., BJORKLUND, F. (2007), "Social intuitionists answer six questions about moral psychology". *In* W. Sinnott-Armstrong (ed.) *Moral Psychology*, vol. 2, Massachusetts, The MIT Press, pp. 181-254.

JACOBSON, D. (2012), "Moral dumbfounding and moral stupefaction". *In* M. Timmons (ed.) *Oxford Studies in Normative Ethics*, 2, pp. 289-315.

KAHNEMAN, D. (2012), *Thinking, Fast and Slow*, London, Penguin.

KANT, I. (1785/1964), *Groundwork of the Metaphysics of Morals*, New York, Harper & Row.

KNIGHT, J., JONSON, J. (1994), "Aggregation and deliberation: on the possibility of democratic legitimacy". *Political Theory*, 22 (2), pp. 277-296.

KUNDA, Z. (1990), "The case for motivated reasoning". *Psychological Bulletin*, 108 (3), pp. 480-498.

LAKOFF, G. (2009), *The Political Mind. A Cognitive Scientist Guide to your Brain and its Politics*, New York, Penguin Books.

LANDEMORE, H., MERCIER, H. (2012), "Talking it out with others vs. deliberation within and the law of group polarization: some implications of the argumentative theory of reasoning for deliberative democracy". *Análise Social*, 205, XLVII (4.º), pp. 910-934.

LE DOUX, J. (1998), *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York, Simon and Schuster.

LEVI, N. (2007), *Neuroethics*, Cambridge, Cambridge University Press.

LIAO, S. M. (2011), "Bias and reasoning: Haidt's theory of moral judgement". *In* T. Brooks (ed.). *New Waves in Ethics*, Basingstoke, Palgrave Macmillan, pp. 108-127.

LORD, C. G., ROSS, L., LEPPER, M. R. (1979), "Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence". *Journal of Personality and Social Psychology*, 37 (11), pp. 2098-2109.

MACINTYRE, A. (1981), *After Virtue: a Study in Moral Theory*, Notre Dame, University of Notre Dame Press.

MARKOVITS, E. (2006), "The trouble with being earnest: deliberative democracy and the sincerity norm". *The Journal of Political Philosophy*, 14 (3), pp. 249-269.

MARTÍ, J. L. (2006), "The epistemic conception of deliberative democracy defended: reasons, rightness and equal political autonomy". *In* S. Besson, J. L. Martí (eds.). *Deliberative Democracy and its Discontents*, Aldershot, Ashgate, pp. 27-56.

MERCIER, H. (2011), "What good is moral reasoning?". *Mind & Society*, 10 (2), pp. 131-148.

MERCIER, H. (2012), "Some clarifications about the Argumentative Theory of Reasoning: a reply to Santibáñez Yáñez". *Informal Logic*, 32 (2), pp. 259-268.

MERCIER, H., LANDEMORE, H. (2012), "Reasoning is for arguing: understanding the success and failures of deliberation". *Political Psychology*, 33 (2), pp. 243-258.

MERCIER, H., SPERBER, D. (2011), "Why do humans reason? Arguments for an argumentative theory". *Behavioral and Brain Sciences*, 34, pp. 57-74.

MONIN, B., PIZARRO, D., BEER, J.S. (2007), "Deciding versus reacting: conceptions of moral judgement and the reason-affect debate". *Review of General Psychology*, 11 (2), pp. 99-111.

NINO, C.S. (1996). *The Constitution of Deliberative Democracy*, London, Yale University Press.

PAXTON, J.M., GREENE, J. (2010), "Moral reasoning: hints and allegations". *Topics in Cognitive Science*, 2 (3), pp. 511-527.

PELLIZZONI, L. (2001), "The myth of the best argument: power, deliberation and reason". *British Journal of Sociology*, 52 (1), pp. 59-86.

PÉREZ ZAFRILLA, P.J. (2009). *Democracia deliberativa. Razón pública y razones no públicas en la perspectiva de John Rawls*, València, Universitat de València.

PIZARRO, D.A., BLOOM, P. (2003), "The intelligence of moral intuitions: comment on Haidt (2001)". *Psychological Review*, 110 (1), pp. 193-196.

RAWLS, J. (1993), *Political Liberalism,* New York, Columbia University Press.

RICHARDSON, H. (2002), *Democratic Autonomy. Public Reasoning about the ends of Policy*, Oxford, Oxford University Press.

SALTZSTEIN, H.D., KASACHKOFF, T. (2004), "Haidt's moral intuitionist theory: a psychological and philosophical critique". *General Review of Psychology*, 8 (4), pp. 273-282.

SANDERS, L. (1997). "Against deliberation". *Political Theory*, 25 (3), pp. 347-376.

SCHKADE, D., SUNSTEIN, C.R., HASTIE, R. (2010), "When deliberation produces extremism". *Critical Review*, 22 (2-3), pp. 227-252.

SPERBER, D., *et al*. (2010), "Epistemic vigilance". *Mind & Language*, 25 (4), pp. 359-393.

SPERBER, D., BAUMARD, N. (2012), "Moral reputation: an evolutionary and cognitive perspective". *Mind & Language*, 27 (5), pp. 495-518.

TAYLOR, C. (1989), *Sources of the Self: the Making of the Modern Identity*. Cambridge, Cambridge University Press.

THOMPSON, D. (2008), "Deliberative democratic theory and practical political science". *Annual Review of Political Science*, 11, pp. 497-520.

TRIVERS, R. L. (1971), "The evolution of reciprocal altruism". *The Quarterly Review of Biology*, 46 (1), pp. 35-57.

WEITHMAN, P. (2002), *Religion and the Obligations of Citizenship*, Cambridge, Cambridge University Press.

WESTEN, D. (2007), *The Political Brain. The Role of Emotion in Deciding the Fate of the Nation*, New York, Public Affairs.

Received 24-09-2015. Accepted for publication 27-04-2016.

PÉREZ ZAFRILLA, P. (2016), "Is deliberative democracy an adaptive political theory? A critical analysis of Hugo Mercier's Argumentative Theory of Reasoning". *Análise Social*, 220, LI (3.º), pp. 544-564.

Pedro Jesús Pérez Zafrilla » p.jesus.perez@uv.es » Universitat de València, Departamento de Filosofía, Facultad de Filosofía y CC EE » Avda. Blasco Ibáñez, 30 — 46010 València, España.