Sponsored since 2011 by the
Italian Society for Analytic Philosophy

ISSN 2037-4445 ©
http://www.rifanalitica.it

Le raisonnement comme moyen de convaincre: Quand l'autorité ne suffit pas?

Dan Sperber (Central European University, Budapest)

[Paris, March, 30th 2015]

Bianca Cepollaro

Let's observe a simplification of *Descartes' conception of reasoning*: it's the capability of distinguishing true from false[1]. According to Descartes, this capability of discerning true from false is what explains human superiority.

Nevertheless, we've got an enigma about reasoning: if *reasoning* is the power of *distinguishing* truth from error and this power is what distinguishes us as humans, why don't we agree on what is true?

Let's consider perception for a moment: the fact that human beings are provided with perception generates a sort of convergence. Of course there are some difficulties, but there is at least a general convergence. We don't observe the same thing for reasoning. There are of course cases where there's strong convergence: imagine that people have to describe a picture with 3 zebras and a giraffe and then evaluate the sentence "there are more zebras than giraffes", by choosing one of the following evaluations: certainly false/probably false/probably true/certainly true. We expect to find a very strong convergence on "certainly true"; but consider now the following case: there are 22 farmers in the village; none of them has more than 17 cows. People have to evaluate the sentence "At least two farmers have the same number of cows", choosing among: certainly false/probably false/probably true/certainly true. The correct answer is "certainly true": yet, only the 30% of people gives the right answer.

Descartes answered to similar concerns saying that the divergence in our opinions depends on the fact that we don't consider the same things and our thoughts follow different leads. Nevertheless, many experiments show that people just reason poorly.

[1] "La puissance de bien juger et distinguer le vrai d'avec le faux, qui est proprement ce qu'on nomme le bon sens ou la raison, est naturellement égale en tous les hommes", Descartes (1637: part I).

Author. Bianca Cepollaro, bianca.cepollaro@gmail.com.

Let's now consider the relationship between perception and reasoning. According to the *Standard view*, besides *perception*, we've got *reasoning*: experimental data show that human reasoning is indeed quite poor. In order to explain that, the standard view distinguishes two mechanisms within reasoning (theory of a double system of inferences):

1. *Intuitive inferences (System 1)*: *rapid heuristics* that work well in most ordinary cases but produce mistakes in non-ordinary situations.

2. *Reasoning (System 2)*: it can check and correct the output of intuitive inferences.

Going back to the farmer case, the intuitive answer would be that it is probably true. We can make a more systematic reasoning and say: there are 17 possible categories of farmers: those with 1 cow, those with 2 cows, etc. Since there are 22 elements, and there is no category that has more than 17 cows, there must be at least one category with more than an element: so we correct our intuition and choose "certainly true" rather than just "probably true".

This standard view we just presented proposes indeed a plausible hypothesis. Nevertheless, things don't work this way: people don't use conscious reasoning to check and correct intuitive inferences, but rather to justify them. Besides, people make not only intuition errors, but also reasoning errors. Reasoning is not generally *more reliable* than intuition. If we look at the literature on human reasoning (Evans 1989), a good example is the so-called "*confirmation bias*" (see also Nickerson 1998), a well-known and widely accepted notion of *inferential error*.

Let's consider now another hypothesis, that puts into question the idea that there is a general system of reasoning. Maybe there are a lot of specialised mechanisms, i.e. *modules*: intuitive inferences are carried out not by a general mechanism of intuition but by many modules, just as perception.

Let's consider now the hypothesis according to which homologous inputs activating different modules may get different interpretations. This hypothesis is compatible with the following result. Consider the farmer case again, in a slightly different version: there are 22 pupil in the class; each got a score between 1 and 17 in the test; people have to evaluate the sentence "There are at least 2 pupils who got the same score" choosing among these options: certainly false/probably false/probably true/certainly true. Even if the problem is the very same as the cows one, people are better at this one (and it's a pretty robust result). Logically, it's the very same thing. This is compatible with the idea that homologous inputs activating different modules may get different interpretations.

Let's now consider the relationship between reason and reasoning: these two things are usually analysed separately (for example, people working on practical reason don't take reasoning into consideration and vice versa). Consider the following case. Night of November 3rd 2013 at Dearborn Heights: Theodore Wafer is awaken during the night by someone; he takes his gun and he kills the person at his door, namely Renisha McBride.

Let's consider how people have been talking about Theodore Wafer's reasons. The *defence* said that he was scared and he was defending himself; the *prosecution* said that the fear was unreasonable: he would not have opened the door if he was scared, he would have just called the police. We observe that the *reasons* that are discussed often have two functions: *explication* (why Theodor did what he did) or *justification* (were his reasons "good" reasons to act?). The motivating reasons have to be seen as justifying by the agent. On the other hand, justifying reasons have to be such that they could motivate the agent.

Notice that this discussion also interacts with the notion of *moral luck*: if the person shot turned indeed out to be a dangerous, armed criminal, T.W. reasons would probably have been judged as good enough.

Let's now take stock and consider a *first approximation of the notion of "reason"*: reasons are the combinations of actual or potential beliefs and desires that to some extent justify accepting some further belief of making some decision (and carrying it out). Beliefs and desires are about some state of affairs. Reason are for some mental representation. Reasons are defined in relation to what they are reasons for; in other words, reasons can be stronger of weaker, better or worse.

A related question is: *do we know our reasons*? We said we have two kinds of reasons: to explain or to justify; the explication-reason is more fundamental: first you explain, then maybe you justify.

Was T.W. conscious of his reasons when he acted? Are we generally conscious? Do we have unconscious reasons that we can then introspect? About this question, see: Nisbett R. E. and Wilson, T. D. (1977).

Another related study is Hall L., Johansson P., Strandberg T. (2012) about choice blindness and attitude reversals on a *self-transforming survey*: subjects had to answer some questions, then they do something else, and after five minutes they have to justify their previous answer. Some people are indeed presented with the answer they provided, some others are presented with an opposite answer, presented as if it was their original one. The interesting result is that the majority of people presented with an opposite answer construe coherent arguments supporting the opposite of their original position, i.e. they don't notice that the one they are presented with is not the answer they gave, rather the opposite. If we had reasons before we give an answer this should not happen: this happens precisely because we form reasons after we choose our answer.

So, *what functions reasons serve*? Primarily, a double social function: *evaluation* (justification or criticism) and *commitment*: to indicate a certain norm, to motivate. We take a responsibility and commit ourselves to behave in a certain way in the future. We can also wonder if *reasons ever guide us*. They may help guiding our actions when we factor in their reputational benefits and costs. I can give up something in order to get something I can justify. Consider the following study: people have to choose one piece of chocolate that comes in two possible shapes: a little heart and a cockroach. People tend to choose the heart, even if the cockroach one is a bigger piece of chocolate. One might conclude that to the extent that people are guided by reasons (for reputational concerns), reasons may help predict their beliefs and actions.

Our intuitions about reasons concern the reasons-beliefs pair and the reasons-decisions pair. These intuitions have both normative and descriptive aspects. The normative aspect of our intuition about reasons is essential to evaluate, justify or criticize beliefs and actions of the people who hold those beliefs etc. In general our intuitions are justified (for example, when the sound of steps is growing louder, we intuitively infer someone is getting nearer. This fact is a good reason to believe that someone is getting nearer). So we can say that reasons fulfil their function in communication in explaining, justifying, criticizing beliefs or behaviours of other people or of oneself.

Let's now consider *reasons in reasoning*: if reasons can justify part of our present beliefs and decisions, they can also be used to convince others to adopt the beliefs they justify (if the circumstances are relevantly similar) or to make the same decision. In other words, the main function of reasoning is to produce reasons to convince others and to evaluate the reasons others produce to convince us. In this sense, we can consider *trust as a reason to believe*: trusting a source is generally a good reason to believe what it communicates. Being trusted is generally sufficient to persuade one's audience. Indeed much of human communication is made possible by trust. Obviously, *trust is common but not universal* nor automatic: humans exercise *epistemic vigilance*. Now, *when trust is not enough*, the communicator may fail to communicate what she intended; a sufficient *authority* is required. What becomes interesting is *the use of reasons to overcome the trust bottleneck*: we may accept information from a source that we do not trust sufficiently if she provides reasons for this acceptance. *Providing reasons* is, for the communicator, a means to convince a reluctant audience. *Evaluating reasons* is for the audience a means to acquire information from an insufficiently trusted source. In reasoning understood this way, the production of reasons should be aimed to persuasion. It should focus on reasons in favour of the conclusion the reasons-producer wants her audience to accept. We also expect that reasoning to persuade should have a confirmation bias.

We can imagine a sort of division of cognitive labour: each group member looked for arguments supporting her perspective and attacking the point of view of other group members. In the end, each opinion is thoroughly evaluated. When people who disagree but share an interest in getting at a good solution argue, reasoning should produce good results. *Wason* selection task (see Moshman and Geil (1998)) shows very interesting results that concern how we perform alone and in groups: 18% correct individual solution; 80% group correct answers. This can be interesting in fields such as justice, educations, research, etc. Note that we are not concerned with the well-being of the group itself, rather with the good of the individuals.

Let's now assess the question about *when collective reasoning does not work*: reasoning is one of many cognitive mechanisms, it's not the only one. Individual reasoning often fails because the confirmation bias is not held in check and indeed people can reinforce their false belief: *group polarization may lead to group fanaticism*.

To sum up, here are the two *conclusions* that can be drawn from this discussion:

1. The main function of reasoning is social.

2. The cognitive and social aspects of reasoning can only be understood together.