



**HAL**  
open science

## A Dialectical Account of Thought Experiments

Jean-Yves Goffi, Sophie Roux

► **To cite this version:**

Jean-Yves Goffi, Sophie Roux. A Dialectical Account of Thought Experiments. Stuart, Michael T.; Fe-  
hige, Yiftach; Brown, James Robert. The Routledge Companion to Thought Experiments, Routledge  
Taylor and Francis Group, pp.439-453, 2018, 9780415735087, 978-0415735087. hal-02052836

**HAL Id: hal-02052836**

**<https://hal.science/hal-02052836v1>**

Submitted on 17 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## A DIALECTICAL ACCOUNT OF THOUGHT EXPERIMENTS

Jean-Yves Goffi (PPL, Université Grenoble Alpes) and Sophie Roux (Mathesis, République des Savoirs, École Normale Supérieure)

One of the many questions that have been asked about thought experiments is the ontological question concerning what thought experiments are, and more specifically, if they are some sort of experiments – perhaps experiments with specific characteristics, but nevertheless experiments (Mach 1905/1926, 149; Sorensen 1992; Buzzoni 2013; Bokulich and Frappier, this volume). After all, if they were not experiments, how could they bring us new genuine knowledge about the world? But, on the other hand, since they are just counterfactual scenarios, how could it make sense to say that they are experiments? Thus the ontological question of what thought experiments are is entangled with the epistemological question of understanding where the knowledge that they bring may come from and what kind of knowledge exactly they can bring. Without dissociating the ontological question and the epistemological question, we will defend in this paper a dialectical account of thought experiments.

In order to defend this dialectical account of thought experiments, we will proceed in the following way:

First, starting from a fairly broad characterization of what thought experiments are in general, we will focus on a tension between two of their characteristics, a tension that puts thought experiments in what we will describe as an ontological state of unstable equilibrium.

Second, having interpreted the epistemological debate between James Robert Brown and John Norton in terms of this ontological unstable equilibrium, we will clarify our position in this debate. To put it bluntly, we agree with Norton when he argues against Brown's Platonist positions, but we do not agree with him when he maintains that thought experiments are only arguments in the sense of valid arguments. According to us, scenarios that thought experiments involve always including some kind of opacity that can not be immediately reduced, thought experiments can not be assimilated to arguments.

Third, we will argue that, to account for the acquisition of knowledge that is at stake in thought experimenting, one should propose a dialectical account of argumentation.

Finally, we will answer some possible objections.

## **1. A Tension in Thought Experiments**

All thought experiments involve exploring the outcomes of counterfactual scenarios in a cognitive context. Without giving a too strict definition of what a thought experiment is, one should note that thought experiments imply three characteristics (Gendler 2004, 1155; Roux 2011, 19–26):

i) First, they involve what we just called a scenario. Such scenarios do not concern entire complete possible worlds; they are rather localized variations, most features of the real world being preserved (Hintikka and Hintikka 1989, 75; Häggqvist, 1996, 146–147). They involve concrete particulars that are usually weaved in short narratives. Some of these particulars are parts of the argument, but some of them are only a psychological crutch for our imagination. For example, if Gyges' thought experiment (see Becker, this volume) can make us think about our motivations when we respect moral norms, it is because there is a device making humans and their actions invisible; whatever this device is, it constitutes so to speak the very bones of this thought experiment, what structures it, and in that sense it is a particular which is essential to the thought experiment. But when Plato tells us that Gyges was a shepherd, who was feeding his flock, when there was a storm, which obliged him to find a shelter in a cave, etc., such particulars do not fundamentally change the thought experiment, they only help us to imagine how it might be that a shepherd comes to possess a magic ring, and in that sense they are inessential, although they help us to consider a story that is not true as relevant and give us reasons to continue with the thought experiment. The distinction between these two kinds of particulars is not always easy to make in practice – to take another example in ethics, many variations were introduced in the Trolley thought experiment initially proposed by Philippa Foot (see Brun, this volume), the outcome of these variations being that what at first sight may have appeared to be inessential particulars turned out to be relevant for our moral judgments. But scenarios are crucial for thought experiments. As we will argue in the second section of this chapter, scenarios involve some kind of opacity and this is the reason why thoughts experiments cannot be reduced to arguments.

ii) The second characteristic is that the scenarios at stake in thought experiments are counterfactual. As opposed to a real experiment, a thought experiment does not have to take place in reality; rather, it seems that we reach its result merely by thinking, a thought experiment being literally an experiment that happens in thought. Here again, some distinctions can be introduced, because there are different degrees of counterfactuality between zero counterfactuality, which corresponds exactly to reality, and maximal counterfactuality, which deals with metaphysical and logical impossibilities. A weaker counterfactuality might concern thought experiments that are physically possible and that could have been produced in reality considering human ability and technical means available at the time they were proposed. However, they were not carried out for whatever reason: this is for example what we see in Galileo's two strapped bodies, the negative part of which can be described as a *reductio ad absurdum* of the Aristotelian thesis that two bodies of the same material descend in the same medium with speeds proportional to their weights. A stronger counterfactuality involves scenarios which not only did not happen, but which, given our human capacity to intervene in the world and the technical means available at the time they were proposed, could not happen. This was for example the case of the EPR thought experiment, at least when it was first proposed in 1935. A stronger counterfactuality still is at stake with Dawkin's thought experiment of rewinding time and playing back history to see if evolution would produce the same results, because it does not depend from technical means (that keep evolving) but make intervene our very concept of time. In the present chapter, we are not interested in the counterfactual aspect of thought experiments.

iii) Third and last, the scenario is framed with a well-defined cognitive intention in a determined cognitive context. If considered independently of their context, Gyges' adventures could belong to a storybook, to a science-fiction novel or to fantasy literature. But they do not, and this is because Plato had a well-determined cognitive intention concerning ethics in view. He wanted his readers to ask themselves how they would behave if they could perform immoral actions without being discovered and, consequently, without being punished. In a word, Gyges' thought experiment plays the role of an argument for Plato, an argument being here, quite generally, what intervenes in a determined cognitive context to make an interlocutor change her mind. Again, distinctions could be introduced. In the case of Gyges' thought experiment, Plato made the reader test her capacity to act according to ethical norms even if her bad behaviour would not be punished. But intervening as an argument can amount to other moves: it can for example amount to introducing a conceptual distinction, to refuting a usually admitted thesis, to testing a declaration by highlighting a paradox, or even proving a previously unknown result (Sorensen 1992, 135–152). In this chapter, we do not want to dwell on these distinctions, but we are definitively concerned with the notion of an argument. We would like to clarify which notion of an argument is at stake when it is said that a thought experiment intervenes as an argument.

That thought experiments present these three characteristics (involving a scenario, being counterfactual, playing the role of an argument) is in general fairly broadly accepted, at least in the recent philosophical literature (see Bokulich and Frappier, this volume, for more on the identify conditions of thought experiments). But, perhaps because of the stress that was put on the counterfactual aspect of thought experiments, it has not been noticed that the first

characteristic and the last one imply that thought experiments are, so to speak, in a state of unstable equilibrium between non-propositional knowledge and propositional knowledge:

On the one hand, thought experiments rely on scenarios in which non-propositional knowledge intervenes. In a former paper, we were led to minimise the importance of the scenario (Goffi and Roux 2011). The point we would like to make in this chapter is that even those particulars that are essential to the scenario are not sufficient to make explicit all the propositions that would be necessary to make a complete argument from a logical point of view. Unlike logical arguments, thought experiments involve scenarios in which non-propositional knowledge intervenes, whether it is called implicit knowledge, tacit knowledge or background knowledge (Gendler 1998, 404–408; Atkinson and Peijnenburg 2004, 121–125).

But, on the other hand, thought experiments are framed with well-determined cognitive intentions and they are used in determined cognitive contexts in which propositional knowledge is important. If the inventor of a thought experiment did not specify why she introduced this or that thought experiment, we could take her thought experiment for a pleasant story to be read before sleeping. Thus, thought experiments are connected to propositional knowledge; the implicit knowledge, tacit knowledge or background knowledge that intervene in the scenarios that they propose is supposed to play the role of premises with regard to a conclusion.

In other words, there is some tension between the non-propositional scenario that a thought experiment involves and the propositional knowledge of the determined cognitive context in which the thought experiment operates. To gain a better grasp of this idea, we can bring experiments and thought experiments closer together. An experiment may intervene in an argumentative context, but is not an argument, simply because it happens that we do experiments without arguing; similarly, a thought experiment intervenes in an argumentative context, but is not an argument for the same reason: it is rather a process that supplies premises for an argument, or that consists in unfolding an initial scenario (Sorensen 1992, 214, 230–240; Häggqvist 1996, 87; Bishop 1999; Häggqvist 2009, 61; Imbert and Skaf 2012). However, the difference between the experiment and the argument that this experiment supports is manifest when one is dealing with a real experiment – even if the experiment is presented in an article or in a book, one assumes that, at some point, a real experiment was actually performed out there and that its outcome could have been something other than what we expected. But the difference between the “experiment” and the argument may be blurred in the case of a thought experiment, because the scenario that the thought experiment describes is not something that happens in reality, but something that unfolds so to speak at the same level as the argument.

To sum up, in the case of a real experiment that intervenes and in the case of a thought experiment, there are two elements, on the one hand a non-propositional element, on the other hand a propositional element. But, while the difference between the two is manifest in the case of a real experiment, it tends to be blurred in the case of thought experiments. This is why thought experiments are, to say it metaphorically, in a state of unstable equilibrium: it is



as if a thought experiment was always on the verge of tilting to one side or the other. What we would like to argue next is that this state of unstable equilibrium may throw some light on the epistemological debate between Brown and Norton. To put it in a nutshell, both of them insist on one characteristic of thought experiments to the detriment of the other, and this, because of their background epistemologies.

## **2. The Unstable Epistemological Equilibrium of Thought Experiments**

As it often happens in contemporary philosophy, the epistemological debate about thought experiments took the form of a choice between two contradictory positions, one of these positions being taken by Brown in his book *The Laboratory of the Mind: Thought Experiments in the Natural Sciences* (first published in 1991, 2<sup>nd</sup> ed. 2011) as well as in a number of other papers where he confronted his views to Norton's, and the other position being taken by Norton, who contributed a paper on Einstein's thought experiments to a volume *Thought Experiments In Science and Philosophy* that was published the same year as Brown's book (Norton 1991), and then published a number of papers where he contrasted his views to Brown's (2004a, 2004b). The debate between them is *prima facie* not a debate on what thought experiments are: Brown and Norton rely on the same set of thought experiments and they would probably agree on a working definition of thought experiments. They rather view their dispute as an epistemological debate concerning the source of the epistemic authority of thought experiments.

For Brown, thought experiments consist literally in observing the outcome of a scenario. In this, thought experiments tilt on the side of non-propositional knowledge; but since this non-propositional knowledge does not build on real experiments concerning the actual world perceived with our ordinary faculties, it has to involve another kind of world, perceived with another kind of faculty. As Brown sees it, in “platonic” thought experiments we are simply “seeing” the universals or “grasping” the relevant law of nature (Brown 1986, 13–12). One would think that we have only metaphors here but, true to Plato, Brown actually thinks that, in contrast with regular experiences where we learn laws empirically, thought experiments give us an intuition of a law of nature, laws being defined as relations among properties (Brown 2004, 33). Häggqvist has nicely summarized: the magic of Brown’s platonic thought experiments consists in their granting “epistemic progress in the absence of new empirical data” (Häggqvist 2007, 50; for a critical appraisal of Brown’s view, see Grundmann in this volume).

Norton denies this kind of epistemic magic. Being an empiricist, he thinks that thought experiments being in thought, they are not real experiments; and that, being not real experiments, they cannot bring us by themselves any kind of new genuine knowledge about the world. Thus, thought experiments have nothing epistemologically remarkable: “thought experiments draw upon what we already know of it, either explicitly or tacitly; they then transform that knowledge by disguised argumentation” (Norton 2004a, 45). In that case, thought experiments tilt on the side of propositional knowledge and standard empiricism is preserved. Norton proposed another justification of his position which he takes to be independent from empiricism: if thought experiments can be used reliably, then they must be

arguments that justify their outcomes, or reconstructible as such arguments (Norton 2004b, 1143). We are not convinced that this justification is independent from Norton's empiricism; it still presupposes the division of labour between experiments, which give us knowledge, and logic, which transforms this knowledge. But, without discussing further, suffice it to refer to the aphorism that Norton himself introduced to summarize his position: "A good thought experiment is a good argument; a bad thought experiment is a bad argument" (Norton 1991, 131).

We seem to face a dilemma, where a choice is to be made between two positions on thought experiments:

1. Either we consider with Brown that there are at least some thought experiments that give us a kind of exceptionally privileged non-propositional knowledge, and we say that their epistemic authority comes from another source of knowledge than our usual sources of knowledge,
2. Or we assert with Norton that thought experiments having no special epistemic authority, they provide only standard propositional knowledge that comes from our usual sources of knowledge, even if, in their case, this standard propositional knowledge appears in a disguised form.

Being completely satisfied neither with Brown's nor with Norton's view, we would like to consider a new option, one that is not associated to any specific epistemological agendas, whether Platonist or empiricist.

In a previous paper, we followed a suggestion of Nicholas Rescher (Rescher 2005) to try to pinpoint the difference between a thought experiment that fails and a thought experiment that succeeds. We described thought experiments as devices aimed at testing the consistency of a set of beliefs when a new belief is introduced as a consequence of the counterfactual situation pictured in the scenario. In a successful thought experience the "experimenter" and her audience are able to reach an agreement on the weakest belief and to eliminate it; a thought experiment fails when it cannot do this (Goffi and Roux 2011, 165-191). Such an analysis amounts to considering that thought experimenters proceed by adjusting beliefs and not by "seeing" with the "mind's eye" laws of nature, or moral norms and values (Brown 2004, 32). In that sense, we are ready to claim with Norton that thought experiments are better reconstructed as arguments. We prefer such a constructivism that conforms to the principle of parsimony to Brown's recourse to Platonic intuitions.

In the present chapter however, we would like to explain why, contrary to Norton, we do not believe that thought experiments are only arguments in disguise. To claim that thought experiments are not arguments in disguise amounts in a way to stating the obvious: if all that is needed to make a complete argument was already given in a thought experiment, this would no longer be a thought experiment, but a logical argument. One can moreover wonder what being an argument in disguise could exactly mean: an argument is explicitly formulated or it

is no argument at all. Last, but not least, if thought experiments were disguised arguments, because of what we call the opacity of the scenarios on which they rely, they would be doomed to be poorer than regular arguments.

The notion of “opacity,” which is essential to our account, can be introduced with a quote from Sören Häggqvist. Having claimed that the connexion between experiments and arguments is that the former supply premises for the latter, Häggqvist explains the differences between regular experiments and thought experiments in the following way: “In the case of laboratory experiments, this [the supplying of premises] is done by causing observers to hold relevant observation statements true (or false) where the causes are the actual physical goings-on in the laboratory. In the case of thought experiments, it is done by causing thought experimenters — whether inventors or audience — to hold relevant non-observational statements true (or false) where the causes are the actual psychological goings-on in the thought experimenter’s heads” (Häggqvist 1996, 87). Without going into details about Häggqvist’s causal theory of belief, we note that he clearly distinguishes two cognitive contexts: in the first context, experimenters deal with actual observations resulting from publicly accepted procedures; in the second context, experimenters deal with actual psychological goings-on triggered by a scenario. In the first case, one has to deal with observer and observations, in the second case one encounters a speaker and its audience. The consequences of the distinction between these two contexts are sweeping: in a thought experiment, much more depends on the narrative than in a regular experiment; more precisely, much more depends on the interpretations that the inventor of the thought experiment and that her interlocutors give of a scenario. If a scenario is opaque in general, it is precisely because it

allows for different interpretations by different interlocutors. It remains to say how. (See Lenhard, this volume, for a different discussion of opacity).

In the first place, there can be different interpretations of the relevance of the scenario with respect to the question asked. In the case of Thomson' violinist, some may believe that the relation between a pregnant woman and her foetus is not adequately captured by the scenario of someone finding herself back to back in bed with an unconscious violinist whose circulatory system has been plugged into hers. According to those dissenters, such a scenario would not have paid enough attention to what is really important in the situation it aimed at describing.

But there can also be different interpretations of the relevance of the scenario with respect to the conclusion that it is supposed to lead to. In the case of Mary the scientist, some may believe that the scenario fails from the start to distinguish between discovering new abilities (knowledge how) and gaining new factual knowledge (knowledge that). According to those, although this scenario clearly points to what is relevant in the situation that it describes, it would lack relevance for the conclusion that is drawn from it.

Finally, a scenario can also be said "opaque" in the sense that some details might be relevant according to the interlocutors, which were not evaluated in the way intended by the inventor of the thought experiment. Bohr famously answered to Einstein's clock-in-the-box that, according to the very theory of relativity discovered by Einstein himself some years earlier, the time as measured by the clock in the box should admit of a minimum level of uncertainty

because of the clock's changing position in the gravitational field of the Earth (for more details on the evolution of this thought experiment, see Bokulich and Frappier in this volume). Thus, although Bohr considered that the clock-in-the box as relevant both for the question and for the conclusion at stake, he judged that Einstein ignored a relevant detail in the scenario, namely the uncertainty of time implied by his very account of general theory of relativity.

This is not to say that a scenario can never be rescued from its opacity. "Opacity" here is not a kind of mystical obscurity that cannot but stay in the darkness. Something opaque is rather something which calls for a process of clarification, a process that in the case of thought experiments will happen, not through "mental reasoning" or "seeing with the mind's eyes," but through actual discussions between interlocutors who disagree. Discussing the possible interpretations and implications of a scenario is thus an important part of the work that the interlocutors perform when they deal with a thought experiment. Sometimes, they will succeed in rescuing the scenario from its initial opacity. But such a rescuing is never ended as long as the thought experiment is a thought experiment; when it is ended, the thought experiment may indeed boil down to a "mere" argument – but then, it is not anymore a thought experiment.

But if, as we argued, scenarios of thought experiments are opaque, how can we describe the acquisition of knowledge that is at stake in them, without hypothesizing some exceptional faculties as Brown and without claiming that thought experiments boil down to arguments in Norton's sense? This is the question that we would like to answer in the next section.

### 3. Steadying the Equilibrium: A Dialectical Account of Thought Experiments

We think that, in order to describe the acquisition of knowledge that is at stake in thought experiments, a strict epistemological view of argumentation is not sufficient. Rather, we will argue that an enlarged view of argumentation should be adopted. We will dub this enlarged view of argumentation “dialectical.” As we will explain in some details, “dialectics” refer here to the domain that Aristotle wanted to establish between rhetoric and science. But speaking of dialectics is also consonant with another idea that we would like to insist on: from an historical point of view, thought experiments never appeared on the private mental scenes of individuals, but in controversial contexts, where a choice between competing theories was to be made. In such controversial contexts, thought experiments were supposed to help the interlocutors to make their choice. But the interpretation of thought experiments appealed in turn controversies, both because of the opacity of their scenarios, that was explored in the previous section, and because of another kind of opacity that we will explore now, the opacity of the background assumptions that were involved in the scenarios.

To introduce this new kind of opacity, let us come back to Brown and Norton for a last time. It is to be noted that both of them begin by assuming that there are some thought experiments that are good thought experiments *per se*, independently of the respective epistemic situations of the thought experimenters. It is only when they want to understand what makes a thought experiment a good thought experiment that some disagreement between them emerges, Brown claiming that its epistemic authority comes from a vision of natural laws, Norton saying that it comes from the disguised argument that it contains, whether this argument is valid in the case of a deductive argument (if one accepts the premises, one is logically bound



to accept the conclusion) or it is at least good enough in the case of an inductive argument (if one accepts the premises, one is logically bound to regard the conclusion as probable).

Our suggestion is rather that we should take into account the epistemic situations of the interlocutors and say that a thought experiment is good when it is successful in rightly convincing one's interlocutor. Even when the opacity of a thought experiment has been reduced, a thought experiment may work or not work. It works when the interlocutors are in such an epistemic situation that they share the same background assumptions and a certain hierarchy of beliefs, so that, when confronted to the inconsistency of their beliefs, they will agree to abandon one belief rather than another. On the contrary, if they do not share the same background assumptions or disagree on the hierarchy of beliefs, the thought experiment does not work. In our previous paper, we insisted on two such situations:

First, it can happen that thought experiments fail to hit on a set of shared assumptions (Goffi and Roux 2011, 184). It is especially striking in what Norton aptly called thought experiment/anti-thought experiments pairs (Norton 2004a, 45-49). As already noted by others, this is for example the case of the two Marys, brilliant scientists who, from birth, were confined to a laboratory, where they had only black and white experiences, but learnt everything about colours and colour-perception. Mary Jackson experiences something new when she leaves her laboratory and sees red for the first time. Mary Dennett however, when she leaves her lab, just exclaims: "Ah! Colour perception is just as I thought it would be! This banana should have been yellow, not blue!" But a similar situation might have been the case also for what has usually been considered as successful thought experiments. Galileo's two

strapped bodies might have been an unsuccessful thought experiment if Salviati had conversed with an Aristotelian who would have refused the assumption that natural speed is mediative, or yet with a physicist who would have been able to take into account the variation of speed with respect to the distance of the earth, or the actual medium in which the bodies fall (Atkinson and Peijnenburg 2004). Thomson's violinist might have been an unsuccessful thought experiment if Thomson's interlocutors had disagreed with her from the start, dispensing with the language of rights and arguing instead that life is a gift from God and that one cannot define a proper stance towards foetal life without being responsive to this property and the personal relationship analogies behind the metaphor. In these cases, Salviati and his opponents, as Thomson and her interlocutors, would have failed to find common ground, exactly as did Jackson and Dennett.

Second, it may happen that two speakers agree on a set of background assumptions, but do not agree on which proposition should be dropped. In the same paper, we gave the example of a kind of counterfactual argument that is to be found in Aristotle's *Politics* (Goffi and Roux 2011, 185–186). Aristotle began asking the question “what would happen if every tool could perform its own work,” but, then, the thought experiment aborts and there is no clear conclusion. We argued that such a failure comes from the fact that Aristotle was dealing with beliefs that were so closely united for him (and probably for his contemporaries) that he could not identify the weakest belief that has to be eliminated, so that all the beliefs had either to stand or to fall together.

Because the same thought experiment may work or not work depending on the background assumptions of the interlocutors and on their hierarchy of beliefs, we think that the epistemic situations of the interlocutors should be taken into account when a thought experiment is evaluated. Of course, this is true not only of thought experiments, but of every argument. But, considering the initial opacity of the scenarios and the role played by background assumptions and the hierarchy of beliefs, thought experiments are more sensible to this context-dependence. In that respect, if we are ready to say that a thought experiments can intervene as an argument in a determined cognitive context to make an interlocutor change her mind, we are not ready to share Norton's notion of an argument, which erases the opacity of the scenario and what could be called the opacity of the background assumptions and of the hierarchy of beliefs. Obviously, the price to pay is to renounce to the idea that thought experiments can pretend to a kind of justification that would be as strong as the justification of a valid deductive or cogent inductive argument. But are there positive references that could substantiate our enlarged view of argumentation?

We think that the tradition of dialectics provide exactly this kind of references. First of all, we must say that we do not claim here a Marxist or Hegelian tradition, where "dialectic" refers to the process by which internal contradictions are transcended, rationally or otherwise. Neither do we think that a dialectical argument contains special logical constants, logical proofs or logical laws different from the usual ones. Referring to the Aristotelian tradition, we consider instead that what sets apart a dialectical argument, above all, is its context. Dialectical arguments have their rightful place in controversial contexts, where the issue is not settled in advance: because the issues at stake are opaque and difficult to clarify, the case allows for

different and initially equally plausible answers. As was aptly shown by Marta Spranzi (2011, 161–172), Aristotle’s dialectical tradition inspired several developments in the field of argumentation theory in the 20<sup>th</sup> century. Without pretending to be exhaustive, we would like to examine three of these developments in order to decide which of them is the fittest to make clear our case about thought experiments.

In their *New Rhetoric*, first published in 1958, Chaïm Perelman and Lucie Olbrechts-Tyteca insisted that their aim was to break with the conception of reason and reasoning that was introduced in the 17<sup>th</sup> century and to rehabilitate the domain that Aristotle used to characterize as the domain of “dialectical argumentation,” that is the domain of verisimilar, likely, plausible and probable arguments. From a strictly logical point of view, such arguments are not valid since they are not constraining proofs that rely on necessity and evidence. But, still, they are much more than the expression of subjective points of view that would imply affects, feelings and instincts (Perelman and Olbrechts-Tyteca 1971, 1-10). However, in practice, the distinction between dialectics and rhetoric tends to be blurred in the *New Rhetoric*. Perelman and Olbrechts-Tyteca rightly point out that dialectical arguments are submitted to an audience and they rightly distinguish between particular audiences (made of *de facto* publics whom the speaker actually addresses in specific situations) and the universal audience (made of a *de jure* public who, in an ideal situation, would rely on logical proofs only or who, from the point of view of God, would be constituted of the totality of beings capable of reason). Still, they waver between two interpretations of the notion of universal audience: on the one hand, they consider that scientists addressing their peers rightly think that they address an universal audience (Perelman and Olbrechts-Tyteca 1971, 30–31); but, on the other hand, they bluntly assert that

“each individual, each culture has its own conception of the universal audience” (Perelman and Olbrechts-Tyteca 1971, 33). This amounts to subscribing to an individual or cultural relativism, according to which there would be as many universal audiences as speakers or as cultures. Even if the notion of universal audience is better understood as a regulative notion, in as much as actual parties in an dialogue are “floating incarnations of this universal audience” (Perelman and Olbrechts-Tyteca 1971, 31), since the genuine aim of argumentation is securing an efficient action on minds with the aid of discourses (Perelman and Olbrechts-Tyteca 1971, 19), it seems that it falls under rhetoric.

That there is more than that in dialectics has been nicely captured by Frans H. Van Eemeren and Rob Grootendorst. They advocated what they call “pragma-dialectics,” where argumentation is understood as an element in a critical discussion aimed at “resolving a difference of opinion” (Van Eemeren and Grootendorst 2004, 52). They insist that resolving a difference of opinion is not the same as settling a dispute: a dispute is settled when the initial difference of opinion has been ended by some external procedure (a vote, a judgment or whatever) even if the various protagonists in the dispute may very well, in *foro interno*, stick to their initial opinion. On the contrary, a difference of opinion is resolved “if a joint conclusion is reached on the acceptability of the standpoints at issue on the basis of a regulated and unimpaired exchange of arguments of criticisms”. Sometimes, a party just surrenders to the other party's argument, throwing in the towel, so to speak and retracting her standpoint (Van Eemeren and Grootendorst 2004, 58; for a similar distinction between resolving, closing, and abandoning a controversy, see for example McMullin 1987, 77–82). More interestingly, it may happen that both standpoints are accepted as various possible opinions about the issue. But Van

Eemeren and Grootendorst also miss something: they do not say anything about the possibility that a new standpoint emerges thanks to the efforts of each party to show the weaknesses of the other party's standpoint and the strength of its own standpoint. In other words, they seem to take the opposing views as ultimate views that are not susceptible of evolving, while genuine controversies are achieved when a new standpoint or conceptual system that supersedes the old standpoints of conceptual systems emerges (Freudenthal 2002).

Thus, to Perelman's and Olbrechts-Tyteca's or to Van Eemeren's and Grootendorst's attempts, we prefer Nicholas Rescher's attempt to understand dialectics in the context of a controversy-oriented approach to the theory of knowledge, where a dialectical argument appears as a process for examining claims to knowledge (Rescher, 1977). Rescher considers that formal disputation as practiced in the Universities during the Middle Ages and ever since is a model for rational controversy. Indeed, in medieval disputations, Aristotle and Thomas Aquinas were supposed to be the material rule of truth. Still, what is interesting here is the formal process through which claims to knowledge are authenticated. A contender have to defend a thesis against counterarguments presented by his opponents, arguments and counterarguments being parts of a "process that probes ever more deeply into the grounding of the proponent's thesis, developing in subtle and comprehensive detail the structure or rational support which he envisages for his focal contention" (1977, 5). In other words, in such a disputation, opposing views are not to be taken as *ultimate* (in contrast to what happens in the pragma-dialectics model) but as various *knowledge claims* about a controversial issue (in contrast to what happens in the new rhetoric model). In a dialectical argumentation, knowledge claims may stay plausible, but not yet secured. They are plausible in the sense that they have at first sight credentials that have to be

examined. After a careful analysis of these credentials, they may be deemed relevant or irrelevant; but often the verdict will be a verdict of partial relevance: a reasonable doubt remains, because a pertinent objection has been raised. In that case, the analysis will go on. In any event, claims to knowledge are carefully tested, and we say that a powerful tool for carrying out such tests is thought experimenting. Dialectically understood, a thought experiment is thus an argument in a process through which knowledge claims are tested: Do they involve a contradiction? Are they as obvious or commonsensical as they seemed at first sight? Are they based on an ambiguity in their formulation? If this is the case, the party who made the assumption must, at least provisionally, bear the burden of proof and put forward a new argument, a new evidence or a more sophisticated approach of her own thesis.

Let us come back to Thomson's violinist. This thought experiment intends to show that the right to life cannot be understood in such a way that it would give a person the right to use another person's body against her will, even if it is a matter of life and death. As we noted earlier, the experiment does not work if the opponents view life as a gift from God, so that, according to them, the language of rights is not appropriate to settle the issue. But from a dialectical point of view, there is still something to learn even if the experiment does not work. The proponent of the thought experiment will for example remind her opponent that she will encounter many problems if she really intends to defend the idea that life is a gift from God (see for example Pabst Battin, 1995, 26-74). Unsuccessful as it was in convincing the opponent, Thomson's thought experiment would have helped to show where the disagreement exactly lies. Thus, the initial opacity will have been reduced: even if the protagonists fail to reach common ground, they will see the issue more clearly than was the case before. Maybe a

definitive answer will forever stay out of reach, but the proposed answers of the opponents will rest on firmer grounds.

#### **4. Preventing Disequilibrium: An Answer to some Objections**

We would like to end our chapter by briefly answering three objections that might be addressed to our thesis that thought experiments are dialectical arguments.

i) The first objection is the following. If thought experiments are only dialectic arguments, how is it that Norton is able to reconstruct them as logically valid arguments or that we were able to present them as a way to test the consistency of a set of beliefs? Or still, how is it that Sorensen and Häggqvist are able to regiment them in logical models? Let us discuss these questions in the case of Häggqvist, who quite convincingly shows that the important class of thought experiments that refute theories can be enrolled in the following logical model (Häggqvist 1996, 97–102; Häggqvist 2009, 63–68):

C (C being the counterfactual situation described in the scenario) is possible.

T (T being the target theory) implies that C necessarily implies W (W being a statement that, considering T, should be the case in the situation described by C).

However C necessarily implies  $\neg W$ .



Hence, the target theory  $T$  is falsified.

This model is elegant, because it is simple from the logical point of view: it uses only four lines in modal propositional logic. It is fecund as well, because of the consequences that might be drawn from the fact that it is a modalized version of the hypothetico-deductive model. This is not a surprising fact, since it is restricted to thought experiments that test theories (Häggqvist 2009, 63). But this allows Häggqvist to show that such thought experiments are threatened by holism for the same reasons as crucial experiments are (Häggqvist 1996, 107–117). There are namely three ways to escape the conclusion that  $\neg T$ , whether by denying that  $C$  is possible, or by denying that  $T$  implies that  $C$  implies necessarily  $W$ , or, still, by denying that  $C$  necessarily implies  $\neg W$  (Häggqvist 2009, 65–69).

If we admit that a good number of thought experiments are to be enrolled in such a logical model, how can we pretend that thought experiments are dialectical? At stake is the interpretation of what a logical model means. We think that the regimentation of thought experiments in such a logical model does not imply that everything is constraining in thought experiments. On the contrary, as Häggqvist himself notes, such a model helps us to see many reasons why thought experiments may be unsuccessful in certain contexts and to illuminate the disagreement that they can inspire. In the case of real experiments that intervene according to the hypothetico-deductive model, one has controlled and public observations that provide “intersubjectively accessible checkpoints”; in the case of thought experiments neither the statement describing the counterfactual scenario nor the statement describing its outcome ( $C$  necessarily implies  $\neg W$ ) are such “intersubjectively accessible checkpoints” (Häggqvist 1996,

118). In a word, the regimentation of thought experiments in a logical model does not eliminate dialectics, it only helps us in seeing where dialectics intervenes.

ii) The second objection comes from the idea that thought experiments would be more reliable in the sciences than in philosophy. There were indeed philosophers, for example Kathleen Wilkes in her book *Real People*, who defended the idea that an essential distinction between scientific and philosophical thought experiments should be introduced. According to Wilkes, scientific thought experiments describe adequately their background, deal with natural kinds, and introduce impossible assumptions only if they are not relevant for the conclusion, while philosophical thought experiments give an inadequate description of their background, deal with indeterminate terms of ordinary language, and rely on assumptions that are both impossible and directly relevant for the conclusion (Wilkes 1992, 13–47; see also Thagard 2013). In these circumstances, the second objection is the following: when we enlarge the conception of argument, are we not trying to save philosophical thought experiments to the detriment of scientific thought experiments? In other words, are we not downgrading sound scientific arguments to dialectical tricks and throwing the baby out with the bathwater?

Our answer to this objection relies on our position about the relationship between scientific and philosophical thought experiments. To put it in a nutshell, we do not see why there would be an essential difference between thought experiments in the sciences and in the other fields of knowledge (Cooper 2005; Sorensen 1992, 11–15; Atkinson and Peijnenburg 2003, 317–318). But we nevertheless think that there is a *de facto* distinction, which is due to the fact that thought experiments have a different cognitive environment in the sciences than the cognitive

environment that they have, for example, in philosophy. Without attempting to describe in detail the cognitive environment in the sciences, we would like to note that, in the sciences, real experiments are not the only “intersubjectively accessible checkpoints”; there are also the controlled extrapolations that are made from these experiments, laws and principles that often trump all other propositions, mathematical formalisms that help us to avoid incoherence, and even general theories. All these devices form the cognitive environment of thought experiments in the sciences, and they guide our modal intuitions when we have to deal with thought experiments. Thus, it is not necessary to suppose that there is an essential difference between thought experiments in the sciences and in philosophy to explain that they yield different results. Suffice it to say that they are inserted in different cognitive environments. In that respect, claiming that thought experiments are dialectical is only being realist about thought experiments: *per se*, thought experiments are not reliable and do not provide us with indefeasible justified true belief; if it happens that a thought experiment contributes to prove something, it is because of its cognitive environment.

iii) We will be quite brief on the third and last objection, because it is only a consequence to the second objection. Does not the thesis that thought experiments should be conceived as dialectical arguments open the way to relativism? Speaking of “shared beliefs” and of “common background assumptions” may namely evoke certain relativistic theses.

The answer to this objection is simple: we defend a local contextualism as to thought experiences that does not necessarily lead to relativism. First, relativism does not consist in saying that there are shared beliefs in certain contexts, beliefs that are not shared in others; it

is the thesis that shared beliefs, because and only because they are shared, are good beliefs – in other words the thesis that a shared belief is *per se* a good belief – and that there is no epistemological criterion for choosing among opposing and competing shared beliefs. Second, even if we would be relativists with respect to thought experiments, such a local relativism would not necessarily imply a general relativism, that would be extended to the whole of the physical sciences or sciences in general – precisely because, as we said before, there are in the sciences other devices than thought experiments, devices that all together contribute to the fecundity and stability of the sciences.

In this chapter, we first characterised thought experiments and insisted on the heterogeneity between propositional and non-propositional knowledge that such a characterization implies. This helped us to throw some light on the debate between Brown and Norton, the first one insisting on non-propositional knowledge, the second one privileging propositional knowledge. More importantly, we argued that, in order to escape the dilemma that Brown's position and Norton's position constitute, we should renounce the question that was common to them both, namely the question of what justifies the knowledge that would be provided by thought experiments, and rather conceive thought experiments as dialectical arguments that provide defeasible knowledge. Thus, if we are ready to say with Norton that thought experiments are arguments, we have contrary to him a dialectical understanding of what an argument is. And this is what was meant by our broad characterization of an argument as "what intervenes in a cognitive context to make an interlocutor change her mind". By stating "change her mind," we wanted to underscore the pragmatic dimension of our interpretation;

by stating “cognitive context,” we wanted to stress that it is not the case that any context whatsoever goes.

## **Bibliography**

Armstrong, David M. 1983. *What is a Law of Nature?*, Cambridge: Cambridge University Press.

Atkinson, David and Peijnenburg, Jeanne. 2003. “When Are Thought Experiments Poor Ones?,” *Journal for General Philosophy of Science*, 34, 305–322.

Atkinson, David and Peijnenburg, Jeanne. 2004. “Galileo and Prior Philosophy,” *Studies in History and Philosophy of Science*, 35, 115–136.

Bishop, Michael. 1999. “Why Thought Experiments Are Not Arguments,” *Philosophy of Science*, 66, 534–451.

Brown, James Robert. 1986. “Thought Experiments since the Scientific Revolution,” *International Studies in the Philosophy of Science*, 1, 1–15.

Brown, James Robert. 1991a. “Thought Experiments: A Platonic Account,” in Horowitz & Massey (eds.), *Thought Experiments in Science and Philosophy*, 119–128.

Brown, James Robert. 1991b. *The Laboratory of the Mind. Thought Experiments in the Natural Sciences*, London: Routledge.

Brown, James Robert. 2004. “Do Thought Experiments Transcend Empiricism?,” in Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*, 23–43.

Brown, James Robert. 2007. “Counter Thought Experiments,” *Royal Institute of Philosophy Supplement*, 61, 155–177.

Brown, James Robert. 2013. “Thought Experiments” in Curd and Psillos (eds.), *The*

- Routledge Companion to Philosophy of Science*, 324–335.
- Curd, Martin and Psillos, Stathis, eds. 2013. *The Routledge Companion to Philosophy of Science*, London and New York: Routledge.
- El Skaf, Rawad and Cyrille Imbert. 2013. “Unfolding In the Empirical Sciences: Experiments, thought experiments and computer simulations,” *Synthese*, 190, 3451-3474.
- Freudenthal, G. 2002. “Perpetuum Mobile: the Leibniz-Papin Controversy,” *Studies in History and Philosophy of Science Part A*, vol. 33-3, 573-637.
- Gendler, Tamar Szabó. 1998. “Galileo and the Indispensability of Scientific Thought Experiment,” *The British Journal for the Philosophy of Science*, 49, 397–424.
- Gendler, Tamar Szabó. 2004. “Thought Experiments Rethought — and Reperceived,” *The British Journal for the Philosophy of Science*, 71, 1152–1163.
- Goffi, Jean-Yves and Roux, Sophie. “On the Very Idea of a Thought Experiment,” in Ierodiakonou and Roux (eds.), *Thought Experiments in Methodological and Historical Contexts*, 145-191.
- Häggqvist, Sören. 1996. *Thought Experiments in Philosophy*, Stockholm: Almqvist & Wiksell International.
- Häggqvist, Sören. 2007. “The A Priori Thesis: A Critical Assessment,” *Croatian Journal of Philosophy*, 19, 47–61.
- Häggqvist, Sören. 2009. “A Model for Thought Experiment,” *Canadian Journal of Philosophy*, 39, 55–76.
- Hintikka, Jaakko and Hintikka, Merrill. 1989. *The Logic of Epistemology and the Epistemology of Logic*, Kluwer Academic Publishers.
- Hitchcock, Christopher. 2004. *Contemporary Debates in the Philosophy of Science*, Oxford:

Blackwell.

Horowitz, Tamara and Massey, Gerald J. (eds.). 1991. *Thought Experiments in Science and Philosophy*, Savage: Rowman & Littlefield.

Ierodiakonou, Katerina and Roux, Sophie (eds.). 2011. *Thought Experiments in Methodological and Historical Contexts*, Leiden, Boston: Brill.

Kuhn, Thomas. 1977. "A Function for Thought Experiment," (1964) in *The Essential Tension*, Chicago: University of Chicago Press, 240–265.

McMullin, E. 1987. "Scientific controversy and its termination," in Engelhardt, H.T. Jr. Caplan, A.L., *Scientific Controversies. Case Studies in the Resolution and Closure of Disputes in Science and Technology*, Cambridge: Cambridge University Press, 49-91.

Nersessian, Nancy. 1992. "In the Theoretician's Laboratory: Thought Experimenting as Mental Modelling," *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, II, 291–301.

Norton, John D. 1991. "Thought Experiments in Einstein's Work," in Horowitz and Massey (eds.), *Thought Experiments in Science and Philosophy*, 129–148.

Norton, John D. 1996. "Are Thought Experiments just What You Thought?," *Canadian Journal of Philosophy*, 26, 333–366.

Norton, John D. 2004a. "Why Thought Experiments Do Not Transcend Empiricism," in Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*, 44–66.

Norton, John D. 2004b. "On Thought Experiments: Is There More to the Argument?," *Philosophy of Science*, 71, 1139–1151.

Pabst Battin, Margaret. 1995. *Ethical Issues in Suicide*, Englewood Cliffs (New Jersey), Prentice-Hall.

- Perelman, Chaim, and Olbrechts-Tyteca, Lucie. 1971. *The New Rhetoric. A Treatise*. Am. translation by J. Wilkinson and P. Weaver. Notre Dame, London: University of Notre Dame Press.
- Rescher, Nicholas. 1977. *Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge*. Albany: SUNY Press.
- Rescher, Nicholas, 2005. *What If? Thought Experimentation in Philosophy*, New Brunswick and London: Transaction Publishers.
- Sorensen, Roy. 1992. *Thought Experiments*, Oxford: Oxford University Press.
- Spranzi, Marta. 2011. *The Art of Dialectic between Dialogue and Rhetoric: The Aristotelian Tradition*. Amsterdam, Philadelphia: John Benjamins.
- Van Eemeren, Frans H., and Grootendorst, Rob. 2004. *A Systematic Theory of Argumentation*. Cambridge: Cambridge University Press.
- Wilkes, Kathleen. 1988. *Real People: Personal Identity Without Thought Experiments*, Oxford: Clarendon Press.