

*Rawad El Skaf*

# WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?



Rawad El Skaf

# WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?<sup>1</sup>

## Sommaire



- 1 – Introduction
- 2 – The possibility of the scenario of a TE
- 3 – Demons and Clocks
- 4 – How TEs evolve: background to foreground dynamic
- 5 – Conclusion

It is usually claimed that in order to assess a thought experiment (hereafter TE) we should assess the nomological possibility, or realizability in principle, of its scenario. This is undoubtedly true for many TEs, such as Bohr's reply to Einstein's photon box. Nevertheless, in some cases, such as Maxwell's demon, this requirement should be relaxed. Many accounts of TEs fail in this regard. In particular, experimental and some mental model accounts are too strict, since they always require realizability in principle. This paper aims at analysing the notion of possibility at play in the scenarios of scientific TEs, and sheds some new light on their nature and function.

**Keywords:** Thought experiments, nomological possibility, Maxwell's demon, Einstein/Bohr photon box, evolution of thought experiments

## 1 – Introduction

The literature on TEs has grown exponentially these last decades and different epistemological accounts have been proposed. TEs have been characterized as *limiting cases of experiments* (e.g. Sorensen 1992, Laymon 1991, Buzzoni 2010), *mental models* (e.g. Nersessian 1992, Gendler 1998) *logical arguments* (e.g. Norton since 1991) or *Platonic a priori avenues to laws of nature* (e.g. Brown since 1986). To assess these epistemic accounts, we should analyse how they answer the following two sets of questions.

The first set of questions is about epistemic issues: If TEs can teach us something new about the world and our theories, then what kind of “new” knowledge do they produce? How could TEs give us *new* knowledge without any *new* empirical data? Do they justify their conclusions? And how do they achieve this goal (e.g. by mere argumentation, by intuiting laws of nature, by simulative model based reasoning, and so forth)? Noting that TEs seem to be so diverse and to fulfil different epistemological functions, answers to this set of questions, therefore, cannot be given merely on the basis of one or two examples. Most accounts in the literature purport to answer these questions explicitly: they define what TEs are, how they function and the kind of new knowledge we should gain through them.

The second set of questions is about modal issues and concerns, more precisely, the notions of possibility we should expect to find in TEs: since we are dealing with thought, rather than real, experiments, what restrictions should we impose on the scenario of a (successful) scientific TE? What notion of possibility, as opposed to the actuality of real experiments, should we expect to deal with in TEs? Different answers to these questions lead to locate TEs closer either to our theories

or to the world.

The main goal of this paper is to address the second set, while shedding light on the first.<sup>2</sup> The paper is divided into 4 sections. In section 2, I define the relevant concept of possibility in scientific TEs. In section 3, I elaborate on this concept with two case studies: Maxwell's original demon and Einstein's photon box. In section 4, I use the evolution of the discussed case studies in order to identify a general mechanism of evolution for TEs.

## 2 – The possibility of the scenario of a TE

When a TEer performs a TE he/she usually imagines a scenario (e.g., the free fall of two objects in two situations: alone, then linked together as to form a compound). He/she mentally unfolds it (how should these objects fall in the two situations, according to the theories assumed in the scenario) and obtains a result (the compound should fall both faster and slower than the heavier object). Finally, he/she uses this result in order to say something about a specific theory (Aristotle's theory of free fall is false, the speed of a free-falling object is independent of its absolute weight).

It is usually claimed that TEs involve imaginary (hypothetical or even counterfactual) scenarios, but an explicit description of such scenarios is rarely offered. Here are five elements usually present in a scenario of a scientific TE:

- a - *Theories, laws and principles:* Aristotle's theory of free fall, quantum mechanics, Heisenberg's principle, relativity theory, thermodynamics, kinetic theory.

1 - This paper won the SPS Young Researcher Prize 2015-2016 <http://www.sps-philoscience.org/informations/actualite.php?id=47&PHPSESSID=ituvveaul2chhh6jsjssgsijrr1>

2 - Cf. El Skaf (2016), Ph.D dissertation, chapter 7, for a novel non-reductive non-restrictive account of TEs that aims to answer the first set of questions and characterizes TEs as inconsistency revealers and eliminators, sharing a common structure.

b - *Fictive, mechanic and natural processes*: A demon who separates molecules, weight measuring process, radioactive disintegration.

c - *Fictive and natural particulars*: A demon, a cat.

d - *Dynamics of the particulars, specially the fictive ones*: A demon opens and closes a massless door, Einstein's being who accelerates an elevator.

e - *Idealizations*: Massless door, no air resistance.

Since the scenario of a TE involves particulars with their dynamics and processes, we are tempted to ask whether they are possible. Nevertheless, what notion of possibility should we use? Since the scenario delimits a theoretical framework, the relevant notion of possibility should thus be nomological – *i.e.* relating to a theory, law or principle.

However, restricting ourselves to nomological possibility does *not* mean that scenarios of successful scientific TEs should *always* be possible under *actual* theories, laws or principles. Some latitude must be given to the scientist in conceiving his/her scenario. In this section I will start by defining what I mean by a nomologically possible scenario and argue that this notion generalizes a requirement largely found in the literature. I will then reject this requirement with the aid of Maxwell's demon case study, which contains a scenario whose nomological possibility remains *indeterminate*.

## 2.1 Nomologically possible scenarios

I will start with the definition of nomologically possible scenarios:

*A scenario of a scientific TE is nomologically possible if the scientist imagines some particulars (e.g., photons, a cat), their dynamics (e.g., let one photon escape from a box, place the cat in an opaque box), and mechanic or natural processes (e.g. weight measuring, atomic disintegration). These elements are sufficiently described and possible under some specific and actual theory, law or principle, applicable to the scenario.*

These scenarios describe *in principle, or nearly realizable imaginary experiments*. That is, they describe experiments that could be performed in the real world – modulo, of course, some specific and necessary idealizations. Precisely these idealizations can be questioned by the scientific community (above all by the defender of the criticized theory or by the opponent of a newly postulated theory, law or principle). In fact, the choice of the particulars in some TEs makes these idealizations more or less acceptable. This is especially the case in Galileo's Pisa tower TE, in which Galileo describes objects, of the same heavy material and of a spherical shape, free falling in a rare medium. These particulars allow Galileo to idealize the effect of air resistance, without assuming the existence of a vacuum.

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

It should be noted that in a nomologically possible scenario the scientist could also employ fictive particulars (e.g. Einstein's being) and their dynamics (Einstein's being pulling a rope), yet the resulting fictive process (Einstein's being accelerating an elevator) should be easily *substitutable* with a mechanical or natural process (a rocket accelerating the elevator for example) and the theories that describe it (Newton laws of action and reaction).

We find this notion of possibility in several accounts of TEs, even if it is described differently:

In his 1964 seminal article on TEs, Thomas Kuhn raises and answers the following epistemological question: "How, [...] relying exclusively upon *familiar data*, can a thought experiment lead to new knowledge or to new understanding of nature?" (Kuhn 1964, reprinted in Kuhn 1977, p. 241). Since then, an extensive philosophical literature emerged and different epistemological accounts of TEs have been proposed. Most address Kuhn's puzzle, but they rarely explicitly question its underlying assumption concerning the familiarity of the data.<sup>3</sup> Just before raising his epistemological question, Kuhn assumes that "[g]ranting that every successful thought experiment embodies in its design some prior information about the world, that information is *not itself at issue in the experiment*. On the contrary, if we have to do with a real thought experiment, the empirical data upon which it rests must have been both well known and *generally accepted* before the experiment was even conceived" (*Ibid.*). On the basis of this assumption, Kuhn can give an answer to the following more specific question: "since the situation imagined in a thought experiment clearly may not be arbitrary, to what *conditions of verisimilitude* is it subject?" (*Ibid.*). Such an answer is precisely that "nothing about the imagined situation may be entirely *unfamiliar or strange*" (All emphases added, *Ibid.*, p. 252).

Kathleen Wilkes' (1988) scepticism about TEs in philosophy is mainly based on her demand that scenarios of TEs should be nomologically possible. Wilkes insists that TEs in philosophy are misleading since, unlike TEs in science, they usually involve scenarios that violate known laws of nature. For example, Wilkes rejects Parfit's splitting persons on the ground that "the sum of laws that group us together as human beings (a natural kind category) precludes our splitting into two" (Wilkes 1988, p. 36).

Tamar Szabó Gendler (1998), following Kuhn's account, claims that "[t]o draw a conclusion on the basis of a thought experiment is to make a judgement about what would happen if the particular state of affairs described in some imaginary scenario were *actually to obtain*." (My emphasis, Gendler 1998, p. 398).

In order to account for TEs, Nancy Nersessian (1992) appeals

3 - Humphreys (1993) explicitly addresses and criticizes this familiarity condition but with different arguments from the ones presented in this paper.

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

to mental models of the type posited by some cognitive psychologists for the comprehension of narratives in general. Roughly, her idea is that TEs in science are narratives that function like any other narrative, and as such “thought experimenting is a form of ‘simulative model-based reasoning’. That is, thought experimenters reason by manipulating mental models of the situation depicted in the thought experimental narrative” (Nersessian 1992, pp. 291-2). Nersessian lists five features of TEs characterized as involving mental models. When discussing the second one, she claims that:

“Feature 2: The reader is invited to follow through a sequence of events or processes *as one would in the real world*. That is, even if the situation may seem bizarre or fantastic, such as being in a chest in outer space, there is nothing bizarre in the unfolding. Objects behave as they would in the real world in the presence or absence of gravity. The assumption is that if the experiment could be performed, the chain of events would unfold according to the way things usually take place in the world.” (*Ibid.*, p. 295).

In 1999, Nersessian further claims that “unlike the fictional narrative, however, the context of the scientific thought experiment makes the intention clear to the reader that the situation is one that is to represent a potential real-world situation” (Nersessian 1999, p.19).

Roy Sorensen (1992) claims that a TE is “an experiment [...] that purports to achieve its aim without the benefit of execution” (Sorensen 1992, p. 205). The fifth proposition in his necessity refuter regimented scheme explicitly states: “the content of the thought experiment is a possibility” (*Ibid.*, p. 135). The TEer is thus explicitly committed to this possibility statement.<sup>4</sup>

Ronald Laymon (1991) in his experimental approach also claims that TEs are similar to real experiments: they are “ideal limits of real experimentation” (Laymon 1991, p. 167). He claims that to defend this approach we need to “(1) show that there exists a series of experimental refinements such that real experiments can be made to approach the postulated idealized thought experiment” (*Ibid.*, p. 174).

Finally, Marco Buzzoni (2010) in his transcendental-operational account of TEs claims: “[a]ll thought experiments must be thought of as translatable into real ones, and all real experiments as realisations of thought ones. What thought experiments have over and above real experiments is the mere fact that they exist in a purely hypothetical sphere; what real have over and above thought experiments is the mere fact that they overstep the sphere of the possible, in the experiment’s real execution” (Buzzoni 2010, p. 1).

These restrictions, even if they are stated differently, share a

common underlying intuition best labelled as the Nomological Possibility Thesis (NPT):

NPT: The Scenario of *any* (successful) scientific TE should be nomologically possible.

Thus, Galileo dropping two objects while idealizing air resistance, Schrödinger leaving a cat to his fate facing a deadly mechanism, and even Einstein’s being accelerating an elevator in the absence of gravity are all compatible with NPT.

Nevertheless, there is room for arguing that NPT is too restrictive and even irrelevant for some TEs. In some cases, the TEer conceives a scenario that contains data, processes or particulars, which are “unfamiliar or strange”. The scenario imagined is not explicitly taken to be a “potential real-world situation”, “actually to obtain”, or “as translatable into [a] real one” and so forth, even with all the necessary idealizations. That is, in some TEs, we are presented with *non-nomologically possible* scenarios. In this paper, I shall analyse a scenario whose nomological possibility remains *indeterminate*.

### 2.2 Scenario whose nomological possibility remains indeterminate

As before, let’s start with a definition:

*A scenario of a scientific TE whose nomological possibility remains indeterminate* is a scenario in which under-described particulars and their dynamics are imagined by the scientist, in such a way that they give rise to fictive processes whose nomological possibility remains open.

The nomological possibility remains open in the sense that even with *our best theories* we are not in a position to say whether the processes imagined in such a scenario are possible or not. They are not even presented by the scientist as nomologically possible processes, or even substitutable with nomologically possible processes. Maxwell’s demon TE, analysed in the next section, is a canonical example: Maxwell imagines an under-described, fictive particular (a demon) and its dynamics (the demon opens and closes a massless<sup>5</sup> door) that amounts to a fictive process capable of separating fast from slow molecules.

A fundamental distinction should be noted at this point, namely the distinction between a process and its *result*. Here I am concerned with the former, rather than the latter. More specifically, in the case of Maxwell’s demon, I am assessing the nomological possibility of the processes described in the scenario (*i.e.*, the process of molecular separation), but I am not analysing, at this stage, what this process leads to (*i.e.*, the resulting separation between slow molecules on one side

4 - This is less clear in Sorensen’s account since he analyzes the different kinds of impossibilities in TEs (cf. Sorensen 1992, pp. 200-202) and claims that “we are committed to [...] law contravening thought experiments” (*Ibid.*, p. 201). Nevertheless, this seems in tension with his necessity refuter scheme since in this latter all TEs could be legitimately refuted if the critics consider its content impossible (*Ibid.*, pp. 148-152).

5 - A massless door is nomologically impossible according to our best theories, nevertheless it is an acceptable idealization in Maxwell’s scenario. As we will see, something close to this idealization will be rejected by Feynman.

and fast molecules on the other).

Finally, there are also nomologically impossible scenarios that I will not consider here, for example, because the TEer deliberately assumes some false theory (*e.g.* different laws of motion), or employs some explicitly nomologically impossible processes (*e.g.* Norton's machine that is capable of reversing digestion process<sup>6</sup>).

### 3 – Demons and Clocks

In this section I will analyse two case studies that instantiate both kinds of scenarios defined above, and their “evolutions”: Maxwell's demon and Einstein's photon box. Both TEs employ under-described processes: the demon mode of operation in the former, weighing and time measuring processes in the latter. I will then expose how both TEs have evolved, so to speak, in subsequent versions that better describe these processes. In the Maxwell's demon case, among the many evolutions of this TE, Marian von Smoluchowski (1912) and Richard Feynman (1977) described mechanical processes of molecular separation. In Einstein's photon box, Niels Bohr (1949) proposed a better description, theoretically and technically, of Einstein's scenario. The difference between these two evolutions is that, in the former, the better described scenario was not aimed at refuting Maxwell's TE. While in the latter, Bohr's better described scenario was precisely aimed at refuting Einstein's TE, by refuting its scenario's nomological possibility.

#### 3.1 Maxwell's original demon

James C. Maxwell presented his TE in a letter to Peter G. Tait (1867), searching to “pick a hole in the second law”. It was then published for the first time at the end of *Theory of heat* (1871), in a section called “limitation of the Second Law of Thermodynamics”. One hundred and fifty pages before presenting his demon, Maxwell underlines the difference in nature between the first and the second laws of thermodynamics, with the following argument:

“Admitting heat to be a form of energy, the second law asserts that it is impossible, by the unaided action of natural processes, to transform any part of the heat of a body into mechanical work, except by allowing heat to pass from that body into another at a lower temperature [...].

[S]uppose that a body contains energy in the form of heat, what are the conditions under which this energy or any

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

part of it may be removed from the body? If heat in a body consists in a motion of its parts, and if we were able to distinguish these parts, guide and control their motion by any kind of mechanism, then by arranging our apparatus so as to lay hold of every moving part of the body, we could, by a *suitable train of mechanism*, transfer the energy of the moving parts of the heated body to any other body in the form of ordinary motion. The heated body would thus be rendered cold and all its thermal energy would be converted into the visible motion of some other body.

Now this supposition involves a direct contradiction to the second law of thermodynamics, but is consistent with the first law. *The second law is therefore equivalent to a denial of our power to perform the operation just described, either by a train of mechanism, or by any other method yet discovered.* Hence, if the heat consists in the motion of its parts, the separate parts which move must be so small or so impalpable that we cannot in any way lay hold of them to stop them.” (My emphasis, Maxwell 1871, pp. 153-54)

With this argument, Maxwell is exploring the way we should understand the second law of thermodynamics in the light of the newly formulated molecular theory, according to which “heat in a body consists in a motion of its parts”. In discussing the laws of thermodynamics, Maxwell starts by presenting the first law of thermodynamics (conservation of energy) and Carnot and Thomson's formulation of the second law. He is arguing that the operation he just described involving a “suitable train of mechanism” is not prohibited by the first law and the laws of mechanics (“this supposition [...] is consistent with the first law”). It is, however, in “a direct contradiction” with the second. That is, while the first law could be grounded on the laws of mechanics, the second should “express some essentially non-mechanical aspect of nature. If it is to receive an explanation at the molecular level, that explanation must refer to the smallness of the molecules, or equivalently, to their enormous number” (Klein 1970, reprinted in Leff and Rex 2003, pp. 64-65).

Can we conceive of such a “suitable train of mechanism”? And what does it say about the nature of the second law? Maxwell, with his TE, explains *what* the demon should be capable of accomplishing, without fully describing *how*. His aim is to defend a statistical interpretation of the second law.

Let Maxwell speak for himself:

6 - John Norton (2004) puts forward such a scenario in his critique of experimental accounts. His TE involves a scenario that cannot be obtained in reality, but whose description is made realistic anyway: “For example, a thought experiment quickly establishes that the time reversibility of physical law is not directly expressed in the phenomena. The phenomena manifest a decided unidirectionality in time. To see this, we need only imagine that we locate a familiar process in a device capable of reversing its time order. If the device is large enough to host a banquet, we would find elegantly dressed diners regurgitating the content of their stomachs, chewing it back to pristine morsels and modestly conveying them back to their plates with their forks—a process compatible with the physical laws but otherwise never seen. The thought experiment does not employ a continuous approach to some ideal limit, such as the gradual elimination of friction. Indeed the thought experiment is more effective the more we avoid idealization, that is, the more realistic we make the processes subject to time reversal” (Norton 2004, p. 63, f. 12).

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

“Before I conclude, I wish to direct attention to an aspect of the molecular theory which deserves consideration. One of the best established facts in thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which both the temperature and the pressure are everywhere the same, to produce any inequality of temperature or of pressure without the expenditure of work. This is the second law of thermodynamics, and *it is undoubtedly true as long as we can deal with bodies only in mass*, and have no power of perceiving or handling the separate molecules of which they are made up. But if we conceive *a being whose faculties are so sharpened that he can follow every molecule in its course*, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us. For we have seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

This is only one of the instances in which conclusions which we have drawn from our experience of bodies consisting of an immense number of molecules may be found not to be applicable to the more delicate observations and experiments which we may suppose made by one who can perceive and handle the individual molecules which we deal with only in large masses.

In dealing with masses of matter, while we do not perceive the individual molecules, we are compelled to adopt what I have described as the statistical method of calculation, and to abandon the strict dynamical method, in which we follow every motion by the calculus.” (My emphasis, Maxwell, 1871, pp. 308-309)

Maxwell aims at limiting the scope of the second law, as the title of his section suggests, with the help of his demon. In order to better understand what he means we have to look at the historical context. The classical phenomenological second law of thermodynamics was formulated without reference to the microscopic constitution of gases. In fact, it was originally

formulated in the context of the caloric theory of heat, which defines heat as a moving substance. With the development of the kinetic theory, heat became defined in terms of molecular motion. It was argued by Clausius, and early Boltzmann, that the second law remains universally valid, applicable to any scale, and should be grounded on a dynamical method that traces each molecule in its course.

Maxwell thought differently<sup>7</sup> and used his demon to limit the scope of the second law of thermodynamics to aggregates of molecules, “which we deal with only in large masses”. That is, the second law has a statistical nature and should be limited to the macroscopic domain. It only applies to the observations made by someone who cannot perceive and manipulate individual molecules. If this capacity to perceive and manipulate individual molecules turns out to be nomologically possible or not, remains open for Maxwell, as well as for us. He is arguing that, if the second law is to be grounded on molecular motion, it should reflect the treatment of molecules in bulk. This is what Maxwell meant by “statistical”<sup>8</sup> in his conclusion that we should adopt the “statistical method of calculation, and to abandon the strict dynamical method, in which we follow every motion by the calculus”.

Note that Maxwell’s conclusion is not the *rejection* of the second law, but rather the claim that it should be understood in statistical terms. This seems to be grounded on the idea that the nomological possibility of the demon is open: if the demon was assumed to be realizable in principle, then Maxwell should have, unjustifiably, concluded that the second law is simply false. Let us contrast Maxwell’s conclusion with a similar attempt “to pick a hole in the second law”. According to Daub (1970), Loschmidt (1876) conceived of a similar “Non-Demon” TE without a finite being to argue against Clausius’ statement of the second law. Loschmidt thought that such a device is nomologically possible and used his TE to undermine Clausius’ statement as a basis for the second law. He even went further and claimed that “the terrifying *nimbus* of the second law, by which it was made to appear as a principle annihilating the total life of the universe, would also be destroyed; [... mankind] would have an inexhaustible supply of transformable heat [into work] at hand in all ages” (Loschmidt 1876, p. 135, quoted from Daub 1970, reprinted in Leff and Rex 2003, p. 52).

On the contrary, Maxwell left the nomological possibility of the demon open, he merely described what the demon should be able to accomplish, *not how*. The demon’s mechanism is under-described in the TE. Indeed, we are not given sufficient details of its mechanism to assess its nomological possibility. What is crucial for the TE to function is the *result* of the

7 - Maxwell used a much lyrical style in writing to Tait about the dynamical interpretation: “But it is rare sport to see those learned Germans [i.e., Boltzmann and Clausius] contending for the priority in the discovery that the second law of [thermodynamics] is the Hamiltonsche Princip. [...] The Hamiltonsche Princip, the while, soars along in a region unvexed by statistical considerations while the German Icarus flap their waxen wings in nephelococcygia, amid those cloudy forms which the ignorance and finitude of human science have invested with the incommunicable attributes of the invisible Queen of Heaven” (quoted from Knott 1911, pp. 115-116).

8 - Cf. Myrvold (2011) for a historical analysis of Maxwell’s statistical conclusion, opposed to probabilistic interpretation of the second law.

demon's action: fast molecules on one side and slow ones on the other. This result is nomologically possible under Maxwell's kinetic theory and also not prohibited by a dynamical interpretation of the second law; nothing in the dynamics of individual molecules prohibits such a result.

To put it differently, the TE aims at showing an *external incoherence* between the newly defended kinetic theory and a dynamical interpretation of the second law. In order to restore coherence, Maxwell defended a statistical interpretation of the second law which limits it to the macroscopic scale, to molecules treated *en masse*, where, by definition, we cannot manipulate individual molecules. For the purpose of the TE, we thus don't need a full description of the demon's mode of action in order to assess its possibility.

Maxwell's demon TE left us with the following dilemma: either the demon turns out to be possible, but then the second law should be completely rejected; or the demon turns out to be impossible, then the second law should not be rejected, but still modified. These two alternatives have been explored by subsequent versions of the TE appeared during the twentieth century.<sup>9</sup> The next section is devoted to the explorations of such evolutions of Maxwell's demon.

### 3.2 Nomological impossibility of mechanical demons<sup>10</sup>

Better described demons were, and still are, the subject of many TEs, computer simulations and even real experiments since the beginning of the twentieth century.<sup>11</sup> Moreover, the aim of most of these attempts is to explore whether the second law should be *completely rejected*.

In the early twentieth century, it had become clear that thermal processes were statistical processes. Empirical violations arose in observable fluctuation phenomena, such as Brownian motion of a pollen grain. The demon TE was then used to investigate better described demons instantiated by purely mechanical devices. The investigation was meant to answer (among other goals) the following questions: can the spontaneous existing microscopic violations of the second law be artificially accumulated to produce macroscopic violations? If not, what is the explanation of this failure? In other words, could there be nomologically possible demons, and if not, why? To answer these questions, different TEs, computer simulations and real experiments<sup>12</sup> were proposed. Here is how two TEs were used.

<sup>9</sup> -Actually, there are even recent analyses. For example, Norton (2013) proposes a general argument for the impossibility of Maxwell's demon based on statistical mechanics, in particular on a violation of Liouville's theorem. By contrast, Hemmo and Shenker (2012) argue for the possibility of the demon, still basing their argumentation on statistical mechanics, in particular on a coarse-grained volume of phase space. Both analyses are beyond the scope of this paper.

<sup>10</sup> - For a substantial historical analysis of the twentieth century demons, from Maxwell to Bennett, cf. Earman and Norton (1998-99) and the references therein. Here I will not consider information theoretic exorcism of Maxwell's demon. They are irrelevant to my purpose and they are rightly criticized in Earman & Norton (1999).

<sup>11</sup> - Cf. Leff and Rex (2003) for an overview of many key papers on the subject.

<sup>12</sup> - Cf. El Skaf and Imbert (2013) for a detailed comparison of these three methods in investigating nomologically possible demons.

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

### 3.2.1 Smoluchowski trapdoor

Marian Von Smoluchowski (1912) argued, through many TEs, that it would be impossible for a mechanical device to manipulate molecules as intended. His now famous idea was to divide the vessel with a trapdoor made asymmetric by a spring. The molecules should then accumulate in one chamber and create a density difference, which can be transformed to work in violation of the second law of thermodynamics (Fig. 1). It is hard to see why the mechanism cannot work as a perpetual motion machine of the second kind producing work for free. Nevertheless, Smoluchowski concluded that no work would be produced.

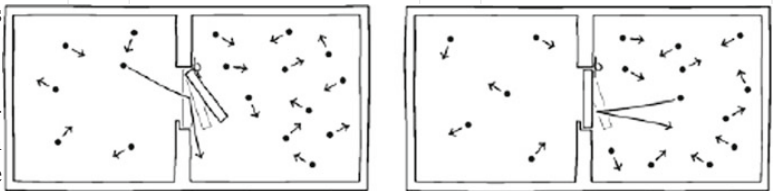


Fig. 1 Smoluchowski trapdoor (picture taken from Norton 2005, p. 401).

Here is how he argued: Since the trapdoor mechanism is in the same isolated container as the gas and needs to be light enough to allow molecules to pass through, it will reach a high thermal energy; thus, the trapdoor will be agitated by its own Brownian motion. Therefore, the trapdoor will open without any collision and, after a period of time, let the passage of molecules in the wrong direction. Thus, this mechanical demon will fail to work as intended. From this device, as well as from different imaginary mechanisms, Smoluchowski concluded that:

“[I]t appears at present that the construction of a perpetual motion machine that produces work continuously is excluded not by purely technical difficulties, *but as a matter of principle*.” (Smoluchowski 1912, p. 1079, quoted from Norton 2013, p. 4439)

Smoluchowski is asserting here that a mechanical demon is nomologically impossible. Which permits him to conclude that:

“Molecular fluctuation phenomena today give us no reason to *overturn completely* the Second Law of thermodynamics, as we have so many other dogmas of physics. They compel us only to a weakened formulation, if we demand universal validity for the laws of thermodynamics.” (My emphasis, Smoluchowski 1912, p. 1079, quoted from

Earman and Norton 1998, p. 447).

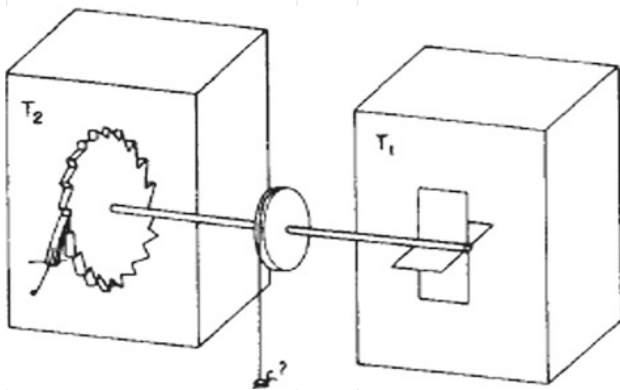
Smoluchowski then proposes the following reformulation of the second law:

“Perhaps an apparently quite minor extension of the wording suffices, in so far as one says: There can be no automatic device that would produce *continuously* usable work at the expense of the lowest temperature. The brief version [of the Second Law] ‘impossibility of a perpetual motion machine of the second kind’ is even sufficient, for one has transferred the difficulty into the explication of the latter concept.” (*Ibid.*)

That is, Smoluchowski is exploring, with his TE, the aforementioned alternatives: either the demon is possible and we should completely reject the second law, or the demon is impossible and we should modify it. Smoluchowski, with his demon TE, argued in favour of the latter.

### 3.2.2 Feynman ratchet and pawl

A very close TE, based on a germane mechanism, “the ratchet and pawl” (Fig. 2) is presented by Richard Feynman (1977). Feynman’s analysis is in the same spirit as Smoluchowski’s. The bombardment of particles in the right container pushes the vane in both directions, but because of the ratchet and pawl in the left container, the mechanism should move in one direction and produce work, in violation of the second law of thermodynamics.



**Fig. 2** Feynman’s ratchet and pawl (Feynman et al. 1977, 46-1, 1).

Nevertheless, Feynman goes on to explain that this result cannot in fact be obtained, he explains what idealizations are not permitted in the imagined mechanism:

“First, our idealized ratchet is as simple as possible, but even so, there is a pawl, and there must be a spring in the pawl. The pawl must return after coming off a tooth, so the spring is necessary.” (*Ibid.*)

13 - The photon box TE was never published by Einstein. Also, his discussions with Bohr during the Solvay conference were not transcribed. It should be noted that there are two different historical interpretations of Einstein’s TE. Cf. Howard (1985, 1990) for a defense of a different version of Einstein’s conclusion, which points at the incompleteness, rather than the inconsistency, of quantum mechanics. Cf. El Skaf (2016) for an analysis of the identity of TEs involving three different photon boxes. In this section I will only present Bohr’s (1949) rendering of this episode.

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

Feynman then explains that he omitted a crucial detail in Fig. 2, which is quite essential. In order to stop the pawl movement after it comes off a tooth, we need to include in the scenario:

“[A]n essential part of the irreversibility of our wheel is a damping or deadening mechanism which stops the bouncing [of the pawl].” (*Ibid.*)

Then, Feynman explains how this mechanism effects the ratchet and pawl:

“When the damping happens, of course, the energy that was in the pawl goes into the wheel and shows up as heat. So, as it turns, the wheel will get hotter and hotter. To make the thing simpler, we can put a gas around the wheel to take up some of the heat. Anyway, let us say the gas keeps rising in temperature, along with the wheel. Will it go on forever? No! The pawl and wheel, both at some temperature  $T$ , also have Brownian motion. This motion is such that, every once in a while, by accident, the pawl lifts itself up and over a tooth just at the moment when the Brownian motion on the vanes is trying to turn the axle backwards. And as things get hotter, this happens more often.” (*Ibid.*, 1-2)

By assuming a canonical distribution function, Feynman then shows by a brief calculation that there is on average no net work done. That is, he concludes that his ratchet and pawl mechanism is nomologically impossible and the second law should not be rejected.

Following this brief historical analysis of the evolution of Maxwell’s demon and the different possibilities of its scenario, I will now turn to Einstein’s photon box TE. This latter case underlines an episode where the nomological possibility of Einstein’s scenario was deemed essential for the success of his TE.

### 3.3 Einstein/Bohr photon box

Einstein, according to Bohr’s (1949) recollection, presented him with the photon box TE during the 6th Solvay conference. Einstein’s aim was to show that the Copenhagen interpretation of quantum mechanics is inconsistent<sup>13</sup>, by attacking Heisenberg’s uncertainty principle.

The uncertainty principle, or uncertainty relations, in quantum mechanics was first enunciated by Werner Heisenberg in 1927. Briefly, it says that the simultaneous measurement of two conjugate variables (such as momentum-position or time-energy) for a given particle, results in a limitation of the accuracy of each of these measures. Namely, the more accurate the measurement of one of the conjugate variables is, the



less accurate it will be for the other. For the energy ( $e$ ) and time ( $t$ ), the better the accuracy of the energy measurement, the less accurate the measurement of the time will be, and vice versa.

With the photon box, Einstein aimed at refuting the time-energy uncertainty relation. According to Bohr, “Einstein proposed the device indicated in [Fig. 3] consisting of a box with a hole in its side, which could be opened or closed by a shutter moved by means of a clock-work within the box” (Bohr, 1949, p. 225). He asks us to start by weighing the box, then open the door for a short time in which a single photon can escape from the box, and finally re-weigh the box. The clock gives us the time of passage of a photon, the balance gives us the difference in the mass of the box before and after the passage of a photon. By using Einstein’s equation  $E = mc^2$  we can calculate the energy difference. This energy difference, in accordance with the principle of conservation of energy, would be the energy of the emitted photon. Therefore, one could in principle simultaneously measure the time of escape of a single photon and its energy, and that with an arbitrary degree of precision. Einstein concluded that this result is “in definite contradiction to the reciprocal indeterminacy of time and energy quantities in quantum mechanics” (*Ibid.*, p. 226).

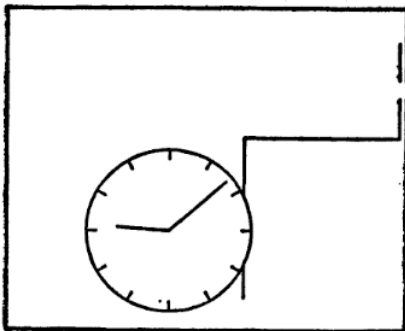


FIG. 7

Fig. 3: Einstein’s photon box according to Bohr (*Ibid.*)

The story according to Bohr ended with his triumph. He found a flaw in Einstein’s scenario the day after that

famous discussion: “After a sleepless night over this argument Bohr rebutted Einstein’s challenge with Einstein’s own general theory of relativity” (Jammer 1974, p. 134). The flaw for Bohr was to be found in the details of the imaginary device of the photon box, in particular the weighing process and its effect on the time measurement by the clock (Fig. 4).

This is how Bohr argued:

“This argument amounted to a serious challenge and gave rise to a thorough examination of the whole problem. At the outcome of the discussion, [...] it became clear, however, that this argument could not be upheld. In fact, in the consideration of the problem, [...] it was *essential* to take into account the *relationship between the rate of a clock and its position in a gravitational field well known from the red-shift* [...].

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

Our discussion concentrated on the *possible application of an apparatus incorporating Einstein’s device* and drawn in Fig. [4] in the same pseudo-realistic style as some of the preceding figures.” (My emphasis, Bohr 1949, p. 226)

Fig. 4: Bohr’s photon box (*Ibid.*, p. 227).

The difference between the two figures is striking, one directly notices the importance attributed by Bohr to the specific details of such a device. Bohr’s analysis concentrates on a better description of the scenario,

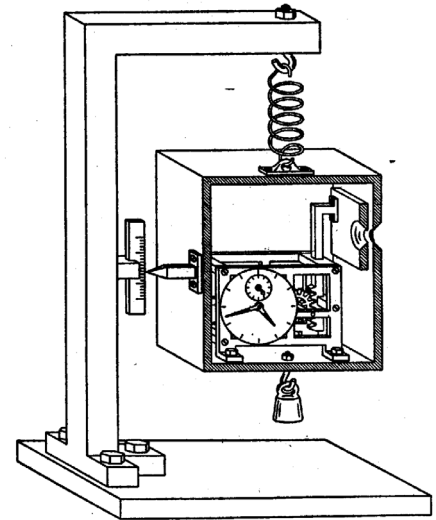


FIG. 8

in particular he fully described how the weighing procedure should be done, with a spring-balance, and the formulas that should describe its behaviour. In particular, Bohr shows that the weighing procedure will require the box to move in a gravitational field. Thus

“[A]ccording to general relativity theory, a clock, when displaced in the direction of the gravitational force [...] will change its rate.” (*Ibid.*)

Bohr is arguing that the accuracy of the time measurement by a clock moving in a gravitational field should be given by Einstein’s own general relativity, and not with a classical space-time theory, as Einstein seemed to suggest.

Which leads Bohr to conclude:

“Consequently, a use of the apparatus as a means of accurately measuring the energy of the photon will prevent us from controlling the moment of its escape.” (*Ibid.*, p. 228)

By fully describing the box weighing mechanism and its effect on the behaviour of the clock reading, since it is now said to move in a gravitational field, arguably what Bohr is doing is to refute the scenario’s nomological possibility under the applicable theory (i.e., general relativity).

Bohr concludes, contrary to Einstein, that we cannot simultaneously and accurately predict the photon’s energy and its time of escape, by doing simultaneous and accurate measurements on the box, since this latter operation is now nomolog-

ically impossible.<sup>14</sup>

## 4 – How TEs evolve: background to foreground dynamic

Both case studies, Maxwell's and Einstein's, and their evolutions, be it in Bohr's reply the day after or Feynman's TE a century later, bring to light a certain dynamic. It consists in the passage from the "background" to the "foreground"<sup>15</sup> of some elements described in the scenario. I take them to be:

- *Background elements*: Processes that are not sufficiently described. The TEer makes reference to such processes when he/she judges that the assessment of their nomological possibility is not necessary for the purpose of the TE.

- *Foreground elements*: Processes that are sufficiently described in a way that we can (and should) assess their nomological possibility under the relevant theories.

The evolution of a TE usually consists in changing some elements of its scenario, in particular by better describing them, theoretically and technically: relevant processes are promoted, so to speak, from the *background* – where their nomological possibility is *assumed* (e.g. Einstein weighing and time measuring processes, Einstein being pulling an elevator in a gravitational free space) or *ignored* (Maxwell's demon mode of operation) – to the *foreground* – where their nomological possibility should and can be *assessed*.

This is what happens in Bohr's reply and in Smoluchowski's and Feynman's TEs. The difference between these two promotions of the relevant processes, from the background to the foreground, is to be found in the purpose of the original scenario, and in the purpose of the evolved one. In the former case, Bohr judged that the processes, sketched by Einstein, should be better described in order to assess his TE. By contrast, in the latter cases, Smoluchowski and Feynman brought the demon mode of operation to the foreground in order to assess if the second law should be completely rejected, and not simply modified or limited. That is, the molecular separation process is better described and promoted to the foreground with the aim of answering different questions and not of refuting Maxwell's TE, or his conclusion.

This dynamic is directly found in Feynman's (1977) analysis. This is how he introduces his TE: "Let us try to invent a device which will violate the Second Law of Thermodynamics [...]. Now let us ask if this is possible." Feynman then tells us "if we just look at [the scenario], we see, *prima facie*, that it

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

seems quite *possible*" and we can conclude that the second law should be completely rejected. Nevertheless, he emphasizes that "*we must look more closely*", and if we do "we see a number of complications" (My emphasis, Feynman et al. 1977, 46-1, 1). That is, in passing the relevant processes from the background to the foreground and assessing their possibility, we should see that the original scenario is nomologically impossible and the second law should not be rejected.

TEs, thus, possess the capacity to evolve, which can be seen as directly linked to the possibility claims involved in their scenarios – specifically, the possibility of the processes described. When presented with a nomologically possible scenario (e.g., Einstein's photon box), we can either challenge the theory (Einstein's conclusion), or the possibility of the processes involved (Bohr's reply). Nevertheless, when the scenario is not presented as nomologically possible, as in Maxwell's original demon, the process' possibility is not, and cannot, be challenged. The TE works even if the scenario presented is not strictly nomologically possible. What matters in Maxwell's demon case is the nomological possibility of the result of such a process – *i.e., fast molecules on one side and slow ones on the other*. This result is nomologically possible under the kinetic theory and nothing in the dynamical interpretation of the second law prohibits such result. This is so, even if the molecular separation process remains in the background, where its nomological possibility is ignored. This suffices to show a tension between thermodynamics and kinetic theory. Maxwell's demon TE, thus, aims at showing an *external incoherence* between these two theories. In order to restore theoretical coherence, Maxwell concludes that the second law of thermodynamics should be statistical.

## 5 – Conclusion

In the introduction, I have separated the questions that each epistemic account of TEs should address into two sets. The first set concerns epistemic issues and addresses what TEs are, how they function, how they justify their conclusions and the nature of the new knowledge they purport to provide. The second set concerns modal issues and addresses the way possibility has to be construed in TEs. I have argued that attention has been paid mostly to the first set. Moreover, when scholars come to the second set, arguably they commit themselves to what I have called the Nomological Possibility Thesis (NPT for short) – *i.e., the claim that a successful TE should involve a nomologically possible scenario*. However, thanks to the Maxwell's demon case, I have shown that NPT is *too restrictive*. Arguably, this is a successful TE which involves a scenario whose nomological possibility *remains indeterminate*. The upshot is that what is at issue in some TEs is

<sup>14</sup> - If all conceivable scenarios were shown to be nomologically impossible by Bohr is a different question that I will not address here. My aim is simply to show that Einstein's description of the relevant processes is insufficient in this case. It should be clear, though, after Bohr's reply, that the burden of proof is now on the critic of the uncertainty principle to furnish a nomologically possible scenario.

<sup>15</sup> - I wish to thank John Norton for his suggestion to use the "background", "foreground" terminology.

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

misrepresented by accounts adopting NPT (e.g., Nersessian's mental model, Gendler's as if actuality, Buzzoni's Transcendental-Operational account and Laymon's or Sorensen's experimental accounts). These accounts are at least not general enough to include TEs, like Maxwell's demon, containing essentially under-described background processes. Moreover, contrary to Kuhn's claim, the data could be "unfamiliar" and "strange". The TEer should in fact be allowed, *in some cases*, to imagine far-fetched scenarios such that their nomological possibility remains *indeterminate*.

The notion of possibility relevant in each TE should thus depend on at least three factors: (i) the function of the TE, (ii) our knowledge of the nomological possibility of the relevant processes, and (iii) the relevance of nomologically possible processes to the TE's conclusion.

Accordingly, the nomological possibility of Maxwell's demon could remain open, since he was not trying to show that we can conceive of an in principle, nearly realizable scenario contradicting the second law. In addition, we don't know about the nomological possibility of all conceivable demons. And even if we did, it is irrelevant to Maxwell's conclusion.

By contrast in the Einstein/Bohr case explicit nomological possibility should be required, since Einstein was trying to show that we can conceive of a near realizable scenario that contradicts Heisenberg's principle. Moreover, we know that clocks run slower in gravitational fields. Finally, the nomological possibility of an accurate clock is relevant to Einstein's conclusion.

Epistemic accounts of TEs should thus permit more latitude to the scientist in conceiving his/her scenario. While some TEs are assessed from what we know about our world and the theories that describe it (e.g., clocks in a gravitational field are inaccurate and are correctly described within relativity theory), others leave the nomological possibility open without assuming a potential real-world situation or its ideal limit (e.g., whether Maxwell believed that our theories, or technology, would someday give us possible, or actual, demons is not an issue relevant to the TE). If we understand these latter as limiting cases of real experiments or as situations that mimic potential real-world processes, we run the risk to weaken a strong TE with misplaced criticism. Here lies the specificity of some TEs compared to other scientific tools, such as computer simulation or real experiments.

My conclusions are based on the analysis of a specific example, namely Maxwell's demon. However, one might deny that this is a TE and thus undermine my conclusions. Ian Hacking (1992), for instance, "resist[s] calling Maxwell's demon part of a thought experiment. It is part of a fantasy. Here I agree with Nersessian, who has remarked to me that it is hard to see what is experimental about the demon. Perhaps

it is only a rhetorical device to reinforce Maxwell's statistical analysis. The demon does not, for me, prove even the *possibility of anything*. [...] But is it a thought experiment? [...] The problem with Maxwell's demon *as an experiment is that you can't conduct it at all*, no matter how much fantasy and idealization you allow yourself" (My emphasis, Hacking 1992, pp. 302-303).

It is true that there is nothing experimental about the demon, since the TE leaves the nomological possibility indeterminate. Hacking concludes that Maxwell's demon is best seen as a "rhetorical device", rather than a TE. The problem is that his claim is clearly based on a specific conception of TEs which flirts with NPT, thus limiting their scope to nearly realizable experiments. Still the question is open: is Maxwell's demon a genuine TE? To answer this question, we need a fine-grained account of what TEs are (*i.e.*, we need to answer the epistemic first set of questions), which is beyond the scope of this paper. Note, however, that Maxwell's demon is readily treated in the literature on TEs as a canonical example of a scientific TE. Hence, there seem to be at least *prima facie* reasons to consider it a TE.

What has been said about the two case studies discussed in this paper can also shed some light on the debate between John D. Norton and James Robert Brown. In addressing the first set of questions, Norton (since 1991) characterizes TEs as deductive or inductive arguments, whereas Brown (since 1986) characterizes some TEs as Platonic vehicles to laws of nature. Although both views could easily account for scenarios whose nomological possibility remains indeterminate, and even for impossible scenarios, they do not explicitly address these issues. Thus, this debate might seem irrelevant to the present purpose. In my opinion this is not the case and some interesting relevant points can be highlighted.

As far as Brown's view is concerned, it is hard to see the role Platonic perception might play in both Maxwell's demon and Einstein/Bohr case. Indeed, Brown does not treat these case studies as Platonic TEs. However, Brown acknowledges, perhaps unwittingly, that Maxwell's demon TE can be seen as an argument *à la* Norton. In his own words: "In one sense, Norton's account fits this example perfectly. We start with the statistical theory and we derive the probabilistic version of the second law; so we have a deductive argument. And the demon, as Norton says, is a 'particular [which is] irrelevant to the generality of the conclusion.' In fact, the demon is utterly unnecessary; we can derive the conclusion without invoking it at all" (Brown 1992, p. 274). It should be clear by now that I disagree with such an analysis, in particular with the irrelevance of the demon and with the idea that the TE was intended to derive a *probabilistic* version of the second law of thermodynamics. This was not the aim of Maxwell's demon TE, simply because the demon violates any version of the second law with a probability equal to a unity. Thus, the

TE would turn out to be ineffective for its (alleged) purpose!

As for Norton, he reduces TEs to “disguised” arguments that involve *irrelevant* and eliminable particulars.<sup>16</sup> Nevertheless, the two case studies offered here suggest that we should give more importance to the nature and function of the elements involved in a TE’s scenario. We have seen how the Einstein/Bohr photon box episode should be analysed in light of the possibility of the processes described in each version of the TE. Precisely these processes prove to be pivotal for both the defence and the rebuttal of the TE. Also, Maxwell’s demon underlines the function of under-described fictive particulars and their dynamics in a TE’s scenario: *they give rise to fictive processes whose nomological possibility remains open*. More precisely, Maxwell’s demon TE raises the following questions: if the argument includes a premise concerning the demon’s action, then what will its truth value be? Will we have a *sound* or *valid* argument for this *good* TE? If the argument does not include the demon’s action, then which premises will it include and will it have the same function or epistemic power as that of the TE? This last question is especially important in the light of the twentieth century, as well as recent, evolutions of the demon’s TE.

#### Acknowledgements

I wish to thank Anouk Barbersousse for her unceasing, and priceless, mentoring; Cyrille Imbert for the many discussions we had during the years; and, Margherita Arcangeli for her comments on a previous version. Finally, I wish to acknowledge my great intellectual debt to John D. Norton. The numerous philosophical discussions we had on thought experiments, as well as his comments on an earlier version of the paper, influenced greatly this final version.

#### RÉFÉRENCES

- BOHR, Niels. 1949. Discussions with Einstein on Epistemological Problems in Atomic Physics. In SCHILPP, Paul Arthur (dir.). *Albert Einstein: Philosopher-Scientist*. Cambridge: Cambridge University Press. 199–241. [Chapitre](#)
- BROWN, James Robert. 1986. Thought Experiments since the Scientific Revolution. *International Studies in the Philosophy of Science*, 1, 1–15. [Article](#).
- BROWN, James Robert. 1992. Why Empiricism Won’t Work. *Proceedings of the Philosophy of Science Association*, 2, 271–279. [Article](#).
- BUZZONI, Marco. 2010. Empirical Thought Experiments: A Transcendental-Operational View. *Epistemologia*, XXXIII, 5–26.
- DAUB, Edward E. 1970. Maxwell’s Demon. *Studies in History and Philosophy of Science*, 1, 213–227. Reprinted in LEFF and REX 2003, 44–58. [Article](#).
- EARMAN, John, NORTON, John D. 1998. Exorcist XIV: The

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

wrath of Maxwell’s demon. Part I. From Maxwell to Szilard. *Studies in the History and Philosophy of Modern Physics*. 29(4), 435–471. [Article](#).

EARMAN, John, NORTON, John.D. 1999. Exorcist XIV: The wrath of Maxwell’s demon. Part II. From Szilard to Landauer. *Studies in the History and Philosophy of Modern Physics*, 30(1), 1–40. [Article](#).

EINSTEIN, Albert, INFELD Leopold. 1938. *The evolution of physics*, Cambridge University Press.

EL SKAF, Rawad, IMBERT, Cyrille. 2013, Unfolding in the empirical sciences: experiments, thought experiments and computer simulations. *Synthese*, 190(16), 3451–3474. [Article](#).

EL SKAF, Rawad, 2016, *La Structure des Expériences de Pensée Scientifiques*. Doctorate dis. Université Paris I Panthéon-Sorbonne.

FEYNMAN, Richard. P., LEIGHTON, Robert B, SANDS, Matthew. 1977. *The Feynman lectures on physics*, vol. I, original edition 1963. Boston: Addison-Wesley Publishing.

GENDLER, Tamar S. 1998. Galileo and the Indispensability of Scientific Thought Experiment. *The British Journal for the Philosophy of Science*, 49(3), 397–424. [Article](#).

HACKING, Ian. 1992. Do Thought Experiments have a Life of Their Own? *Proceedings of the Philosophy of Science Association*, 2, 302–308. [Article](#).

HEMMO, Meir and SHENKER, Orly. 2012, *The Road to Maxwell’s Demon: Conceptual Foundations of Statistical Mechanics*. Cambridge: Cambridge University Press. [Livre](#).

HOWARD, Don. 1985, Einstein on Locality and Separability. *Studies in History and Philosophy of Science*, 16, 171–201. [Article](#)

HOWARD, Don. 1990, “Nicht Sein Kann was Nicht Sein Darf,” or the Prehistory of EPR, 1909–1935: Einstein’s Early Worries about the Quantum Mechanics of Composite Systems. In Miller, Arthur I. (ed.) 1990, *Sixty two years of uncertainty, : Historical, Philosophical, and Physical Inquiries into rhe foundation of Quantum Mechanics*. New York: Plenum Press, pp. 61-111. [Chapitre](#)

HUMPHREYS, Paul. 1993. Seven Theses on Thought Experiments. In EARMAN, John et al. (dir.). *Philosophical Problems of the Internal and External World*. Pittsburgh: University of Pittsburgh Press. 205–227.

JAMMER, Max. 1974. *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in historical perspective*. John Wiley and Sons.

KLEIN, Martin J. 1970. Maxwell, his Demon, and the Second Law of Thermodynamics. *American Scientist*, 58, 84–97. Reprinted in LEFF and REX 2003, 59–72.

KNOTT, Cargill Gilston. 1911. *Life and scientific work of Peter Guthrie Tait*. Cambridge: Cambridge University Press. [Livre](#).

KUHN, Thomas. 1964. A function for thought experiments. Reprinted in *The essential tension*, T. Kuhn, 1977. Chicago: University of Chicago Press. 240–265.

LAYMON, Ronald. 1991. Thought Experiments of Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and as Semantic Domains. In HOROWITZ Tamara, MASSEY Gerald J. (dir.). *Thought Experiments in Science and Philosophy*. Lanham: Rowman & Littlefield. 167–192.

16 - This is Norton’s “elimination thesis. Thought experiments are arguments which contain particulars irrelevant to the generality of the conclusion. Thus any conclusion reached by a good thought experiment will also be demonstrable by an argument which does not contain these particulars and therefore is not a thought experiment.” (Norton 1991, p.131)

- LEFF, Harvey S., REX, Andrew. 2003. *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing*. Bristol and Philadelphia: Institute of Physics Publishing. [Livre](#).
- LOSCHMIDT, Josef J. 1876. Über den Zustand des Wärmegleichgewichtes eines Systemes von Körpern mit Rücksicht auf die Schwerekraft I-IV. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-Wissenschaftliche Klasse*, 73, 128–142.
- MAXWELL, James Clerk. 1871. *Theory of heat*. London: Longmans, Green, and Co. [Livre](#).
- MYRVOLD, Wayne C. 2011. Statistical mechanics and thermodynamics: A Maxwellian view. *Studies in History and Philosophy of Science Part B*, 42 (4), 237–243. [Article](#).
- NERSESSIAN, Nancy. 1992. In the Theoretician's Laboratory: Thought Experimenting as Mental Modeling. *Proceedings of the Philosophy of Science Association*, 2, 291–301. [Article](#).
- NERSESSIAN, Nancy. 1999. Model-Based Reasoning in Conceptual Change. In NERSESSIAN N. et al. (ed.) *Model-Based Reasoning in Scientific Discovery*. Springer US. 5-22. [Chapitre](#).
- NORTON, John D. 1991. Thought Experiments in Einstein's Work. In HOROWITZ Tamara, MASSEY Gerald J. (dir.). *Thought Experiments in Science and Philosophy*. Lanham: Rowman & Littlefield. 129–148.
- NORTON, John D. 2004. Why Thought Experiments Do Not Transcend Empiricism. In HITCHCOCK Christopher (dir.). *Contemporary Debates in the Philosophy of Science*. Oxford: Blackwell, 44–66.
- NORTON, John D. 2005. Eaters of the lotus: Landauer's principle and the return of Maxwell's demon. *Studies in the History and Philosophy of Modern Physics*, 36(2), 375–411. [Article](#)

#### HISTORIQUE

Article initialement soumis le 16 septembre 2016.

Article révisé soumis le 25 novembre 2016.

Article accepté le 29 décembre 2016.

#### SITE WEB DE LA REVUE

[sites.uclouvain.be/latosensu/index.php/latosensu/index](http://sites.uclouvain.be/latosensu/index.php/latosensu/index)

ISSN 2295-8029

DOI [HTTP://DX.DOI.ORG/10.20416/LSRSPS.V4I1.543](http://DX.DOI.ORG/10.20416/LSRSPS.V4I1.543)



SOCIÉTÉ DE PHILOSOPHIE DES SCIENCES (SPS)

École normale supérieure

45, rue d'Ulm

75005 Paris

[www.sps-philoscience.org](http://www.sps-philoscience.org)

## WHAT NOTION OF POSSIBILITY SHOULD WE USE IN ASSESSING SCIENTIFIC THOUGHT EXPERIMENTS?

- NORTON, John D. 2013. All Shook Up: Fluctuations, Maxwell's Demon and the Thermodynamics of Computation. *Entropy*, 15(10), 4432–4483. [Article](#)
- SMOLUCHOWSKI, Marian V. 1912. Experimentell nachweisbare, der üblichen Thermodynamik widersprechende Molekularphänomene. *Physikalische Zeitschrift*, 13, 1069–1080.
- SORENSEN, Roy A. 1992. *Thought Experiments*. Oxford: Oxford University Press. [Livre](#).
- WILKES, Kathleen V. 1988. *Real People: Personal Identity without Thought Experiments*. Oxford: Oxford University Press. [Livre](#).

#### CONTACT ET COORDONNÉES :

Rawad El Skaf  
 IHPST, 13 rue du four  
 75006, Paris  
 France

[rawadskaff@gmail.com](mailto:rawadskaff@gmail.com)

