Evaluating the Cognitive Success of Thought Experiments

Damián Islas Mondragón

Abstract:
Thought experiments are widely used in natural science research. Nonetheless, their reliability to produce cognitive results has been a disputable matter. This study is conducted to present some rules of confirmation for evaluating the cognitive outcome of thought experiments. I begin given an example of a “paradigmatic” thought experiment from Galileo Galilei: the falling bodies. Afterwards, I briefly surveying two different accounts of thought experiments: James R. Brown’s rationalism and John D. Norton’s empiricism. Then, I discuss their positions and I show that none of them may tip the balance towards the rationalism or empiricism they try to defend. Finally, I put forward that the notion of confirmation, connected to the notion of increasing plausibility, can be used to develop some confirmation rules to compare the explanatory power of thought experiments in competition, regardless of their rational or empirical nature in which the discussion of this type of experiment has been engaged in recent years.

Keywords:
Thought Experiments; Background knowledge; Confirmation; Plausibility; Success

Received: 05 August 2017. Reviewed 31 October 2017. Accepted: 30 November 2017. DOI: http://dx.doi.org/10.24117/2526-2270.2017.i3.06

Introduction

Thought experiments are “unexecuted experiments” devised to support or refute a scientific theory – or part of it. Although they are conducted in the mind of scientists, most of them are formulated from certain empirical background knowledge previously accepted. Thought experiments used in the natural sciences have an ancient tradition. For example, Galileo Galilei (1638 [1914]) used them to contrast his theory of free falling bodies against the prevailing theory of his time developed by Aristotle. Albert Einstein et al (1935) set up the famous EPR thought experiment against the completeness of Copenhagen interpretation of quantum mechanics. Recently, Pierre-Marie Robitaille (2014) suggested a thought experiment to refute Kirchhoff law of thermal emission. However, despite their importance for scientific knowledge, their nature and reliability has been a disputable matter.
Until few years ago, we only had three systematic philosophical studies on thought experiments, i.e., the works of Ernest Mach (1906 [1896]), Karl R. Popper (1959 [1934]) and Thomas S. Kuhn (1977 [1964]). According to Kuhn, “The category ‘thought experiment’ is in any case too broad and too vague for epitome” (Kuhn 1964, 241). Nancy Nersessian, whom has written extensively about thought experiments, has recently accepted that: “There is great variety among thought experiments and it would be an impossible task to construct a list of all their salient features” (Nersessian 2007, 147). There is no consensus on the nature of thought experiments. Mélanie Frappier et al puts the things this way “[... ] there is no consensus on the cognitive power of thought experiments, their logical character, the nature of their content, or the proper domains of their application” (Frappier 2013, 1). On the contrary, some authors, leaving aside this problem, think that we do not need a definition of what thought experiments are to assess their cognitive significance for scientific development (see Bunzl 1996; Peijnenburg and Atkinson 2003; Moue et al 2006 and Urbaniak 2012).

The aim of this paper is not, of course, solving this debate; but to present some rules of confirmation for evaluating the cognitive success of thought experiments regardless of the nature of their content. In section 1, I present Galileo’s falling bodies thought experiment as an example of a “paradigmatic” thought experiment. In section 2, I briefly surveying two different accounts of thought experiments: James R. Brown’s “platonic” rationalism and John D. Norton’s “modest” empiricism. In section 3, I show that none of them may tip the balance towards the rationalism or empiricism they try to defend. In section 4, I suggest – following the germinal ideas of Theo Kuipers (2000) – that the idea of “confirmation”, linked to the notion of “increased plausibility”, can be used to develop some confirmation rules to compare the explanatory power of thought experiments in competition, regardless of their rational or empirical nature.

**Galileo’s Falling Bodies**

According to Galileo Galilei (n. 1564 - 1642), Aristotle did not share the belief, already ancient in his time, that vacuum is a prerequisite for the laws of motion. As is well known, Aristotle argued that motion was precisely the phenomenon that made the idea of “vacuum” unsustainable. Galileo said that Aristotle never performed any “real” experiment to show that bodies with different weights travel in the same medium at speeds that are proportional to their weights, as Aristotle erroneously supposed (Galileo 1914 [1638]). Let’s remember that according to Aristotle, if we think about two rocks that fall simultaneously from a height of one hundred cubits, one of which weighs ten times more than the other, we can infer, Aristotle reasoned, that when the heaviest rock reaches the ground, the lightest one would have fallen no more than ten cubits. That is, it would only have covered a tenth of the distance traveled by the heaviest rock.

Galileo assured that he did perform some experiments consisting of simultaneously dropping a cannonball weighing up to two hundred pounds and a musket ball weighing only half a pound from a height of two hundred cubits. However, Galileo asserted (in the voice of his interlocutor Salvati) that it is possible to show – by a thought experiment – that the heavier body does not move faster than the lightest body without executing a “real” experiment. Galileo’s thought experiment is as follows.

Suppose we have two rocks. One is larger than the other. The largest rock moves at a speed of eight cubits while the smaller rock does at a speed of four cubits in free fall. Now

---


3 The cubit was a unit of anthropometric length used in many ancient cultures which corresponds to the distance between the cubit and the end of the open hand.
suppose that we form a system with these two rocks by joining them together with a rope. If both rocks are dropped simultaneously, we can expect that the smaller rock, acting as a kind of drag, slow down the movement of the larger rock, and thus the entire system would fall at a lower velocity than eight cubits, that is, at a lower velocity than the move of the larger rock falling alone. However, we can also expect that the whole system will fall at a greater rate than eight cubits, if the two rocks are considered to constitute a larger unitary system than either of the two isolated rocks. In short, the thought experiment shows us that the fastest rock would be partially retarded by the slowest, and in turn, the slowest rock would be somehow accelerated by the faster one. As we can see, Galileo concluded, this hypothetical situation contradicts the original assumption of Aristotle (Galileo 1914 [1638], 63).

The thought experiment elaborated by Galileo allowed him to infer that when two bodies are in motion, the smaller body attached to the larger one does not “add” its weight to the latter and consequently does not increase its weight as when both bodies are at rest. Galileo concluded that, contrary to Aristotle’s assumption, large and small bodies have the same specific gravity regardless of their weight, so they move at the same speed, reaching the ground at the same time when they are left in freefall (Galileo 1914 [1638], 64-65).

Note that the success of the thought experiment proposed by Galileo lies in the empirical assumption that there is some delay in the free fall of the entire system caused by the small rock. Of course, the way this system moves was not considered by Aristotle. This means that the calculation elaborated by Aristotle and the thought experiment devised by Galileo do not share the same empirical presuppositions although both, as Ana Butkovic says, “mean the same thing by the word ‘faster’” (Butkovic 2007, 65).

Certainly, Aristotle’s concept of “speed” was an essential part of his theory of motion and had important implications for the whole of his physics. But as Kuhn suggested, those implications could never have been challenged solely from empirical observation or logical...
rules in a “world” where all motions were uniform, as in Aristotle’s “world”. The concept itself did not show any logical inconsistency, but a failure to fit, from the subsequent point of view of Galileo’s physics, the “full fine structure of the world to which it was expected to apply” (Kuhn 1964, 258). So, the aim of Galileo’s thought experiment was to reveal an internal contradiction involved in Aristotle’s theory of motion. Galileo used his thought experiment as a theoretical instrument with a specific purpose. As we will see, this use of thought experiments will lead us to develop some confirmation rules to compare the explanatory power of thought experiments, regardless of their rational or empirical nature in which the discussion of this type of experiment has been engaged in recent years.

**Brown’s Rationalism**

There is a considerable amount of philosophical literature around the well-known debate on thought experiments between Brown’s “platonic” rationalism and Norton’s “modest” empiricism. One of the main Brown’s contributions to the topic is his taxonomy of thought experiments, which break into two general kinds i.e. destructive and constructive thought experiments. Constructive thought experiments break into three further kinds i.e. direct (which start with a well-established phenomenon and end with a well-articulated theory), conjectural (which scientists try to establish some phenomenon) and mediative (which are used to derive a conclusion from a well-articulated theory). But there is a small class, the so-called platonic thought experiments, which are simultaneously constructive and destructive, according to Brown. The importance of platonic thought experiments comes from their ability to destroy an existing theory and simultaneously build a new and presumably better one. Let’s briefly see the process.

According to Brown, laws of nature play a crucial metaphysical role in scientific knowledge of the physical world. A law of nature is an “independently existing abstract entity—a thing in its own right that is responsible for physical regularities” (Brown 2011, 199). As an abstract entity, laws of nature supposedly “exist” outside space and time. As claimed by Brown, platonic thought experiments can generate a priori knowledge, this is, knowledge that is not based on new empirical evidence. This a priori knowledge is gained by a kind of perception of the relevant laws of nature with the “mind’s eye”. In Brown’s words:

> Just as the mathematical mind can grasp (some) abstract sets, so the scientific mind can grasp (some of) the abstract entities which are the laws of nature. (Brown 2011, vii)

In other words, since laws of nature “are relations among universals” and these “relations among abstract universals explain observed regularities in the physical world” (Brown 2011, 87), we can presumably know, with the help of platonic thought experiments, the natural world through the a priori access to the abstract realm where these laws of nature supposedly “inhabit”. So, according to Brown, the main cognitive function of platonic thought experiments is producing a priori knowledge of the natural world gained through intuitions.

**Norton’s Empiricism**

Departing from an empiricist point of view, Norton says that thought experiments in natural science are merely picturesque arguments in which scientists “[…] (i) posit hypothetical or counterfactual states of affairs, and (ii) invoke particulars irrelevant to the generality of the conclusion” (Norton 1991, 129, original emphasis). The knowledge produced by thought

---

5 Cammilleri (2015 & 2014); Clatterbuck (2013); Bishop (2012); McAllister (2004); Gendler (2004; 1998); Borsboom et al (2002); Massey (1995); Holton (1993) and Gooding (1992) are some examples.
experiments “[… ] comes from premises introduced explicitly or tacitly into the thought experiment” (Norton 2004b, 1140). This knowledge is then transformed using deductive or inductive inferences to find out about the world without drawing on new information from the world.6

The key point of Norton’s position is that the cognitive outcome of thought experiments is reliable only if the held information presupposed in the premises of its argument form is true and we preserve its truth, or its probability, using valid argumentative process. Note that in Norton’s account, the reliability of thought experiments matches the reliability of their argument form. This thesis is known as the “reliability thesis” (Norton 2004b, 1143) and is the main reason why thought experiments are epistemically unremarkable, this is, they cannot do more than an ordinary argument can do.

Since it is not obvious that all thought experiments are arguments, we must reconstruct them as such. In this sense, Norton asserts that he has not found any thought experiment that cannot be reconstructed as an argument. Some authors have tried to support this idea. For instance, Rafał Urbaniak recently made a plausible reconstruction as an argument of Galileo’s falling bodies thought experiment (Urbaniak 2012). Others, of course, disagree and have tried to show that some thought experiments cannot be reconstructed in such a way (e.g. Bishop 1999; Borsboom et al 2002; Gendler 2004 and Nersessian 2007).

In sum, Norton’s equivalence between thought experiments and arguments provides us with a general criterion of demarcation between good from bad thought experiments, i.e., a good thought experiment is a good argument while a bad thought experiment is a bad argument. So, a thought experiment is epistemically justified insofar as its argument form can justify its conclusion.

**Discussion**

The supposedly “scientific intuition” gained by platonic thought experiments represents a key difference between Brown’s ideas and Norton’s position on the issue. In Brown’s words:

> A way of seeing the difference between Norton and me is to consider, first, real experiments. We would agree (as would most people) that a real experiment carries us from a perception (and some possible background propositions) to a proposition (a statement of the result). I hold that a thought experiment has a similar structure. The only difference is that the perception is not a sense perception but, rather, is an intuition, an instance of seeing with the mind’s eye. (Brown 2004a, 35)

Nevertheless, Brown’s position does not explain the way scientists acquire this kind of “intuition”. Instead of building a reliable epistemological justification for its use; Brown justifies his position appealing to an inference to the best explanation. A debatable position by itself. In Brown’s words: “Readers who find the ontological richness of Platonism distasteful should simply recall that the alternatives are even less palatable” (Brown 2011, 74).

Besides, it seems to me that another central problem of Brown’s rationalism is that platonic thought experiments “are fallible”, as he himself claims (Brown 2011, 42). But if this is the case, then the use of the terms “constructive” and “destructive” is merely tentative, and thus, the alleged essence of platonic thought experiments is straightforwardly uncertain. An epistemological consequence of this uncertainty is that Brown’s stance does not provide any way for assessing the cognitive content of thought experiments and, therefore, neither an accurate procedure for distinguishing successful from unsuccessful thought experiments.

---

6 Note that some thought experiments are not related to the natural world. For example, thought experiments in pure Mathematics or in Ethics.
In relation to the empiricist stance developed by Norton, he asserts that a thought experiment has a justified outcome “if it proceeds from true premises” and that we cannot justify its cognitive outcome if the conclusion comes from (a) a fallacious inference or (b) the utilization of a false assumption (Norton 2004a, 64). But, since Norton avoided to explain how can we know if the explicit or tacit held assumptions presupposed in the premises of the argument form of a thought experiment are indeed true or false, then we neither know if its cognitive outcome is reliable or not. At this point, Norton is quite ambiguous.

Now, although he says that his view is “a consequence of a modest empiricism” that aims to show that thought experiments are not more than “picturesque arguments” and that disguised arguments allows us to develop a “simple empiricist epistemology of thought experiments in the natural sciences” (Norton 1996, 334 and 2004a, 55 emphasis added). At the same time, he contends that the justification of his epistemology of thought experiments is “independent of empiricism” (Norton 2004a, 52). To justify this apparent contradiction, Norton asserts that even if the philosophical stance of thought experiments as arguments is entailed by empiricism (Norton 1996: 335), it is not equivalent to it because:

One would need to place restrictions on the character and relationship of the premises. For example, one would have to assume that we cannot generate conclusions with experiential content unless there are premises with at least as much experiential content [...] In principle, one may hold the argument view without any commitments concerning the origin of the premises used in the argument and their connection with experience. (Norton 1996, 336-337)

In other words, Norton says that thought experiments are arguments that explicit or implicitly holds true or false premises. But I think that asserting that premises may be true or false is trivial. Of course, we know that scientists can justify the outcome of a thought experiment if its premises are true and that they cannot do it if its premises are false (here the triviality). The point is that if somebody aims to uphold an empiricist position (modest of not) on thought experiments, the key point is to make explicit under which cognitive conditions the premises of the argumentative form of thought experiments are true or false.

Finally, Galileo’s thought experiment leads us to think that there is another way to assess thought experiments beyond the rationalist and empiricist positions contended by Brown and Norton respectively. When a thought experiment exhibits some success in achieving certain cognitive goals – in this case pointing out an internal theoretical contradiction of Aristotle’s ideas – we tend to consider this fact as a sufficient evidence to assert that the thought experiment in question is epistemically reliable. That it reaches its cognitive goal. This position can be called the instrumentalist position of thought experiments according to which the main purpose of a thought experiment is to achieve a specific cognitive outcome. In this sense, note that Galileo’s thought experiment does not imply that the two rocks will reach the ground at the same time, as his theory of free fall claims. His thought experiment only aims to reveal an internal contradiction involved in Aristotle’s theory of motion.

Nevertheless, here I will not further explore how can we empirically or rationally assess the cognitive content of thought experiments or if it is sufficient to consider the cognitive outcomes of thought experiments to assess their epistemic reliability or whether their epistemic reliability depends on more than the display of successful cognitive outcomes. These issues are important and show some future lines of research on this topic. Nevertheless, in what follows what I will do is to suggest some rules that can help us to compare the cognitive success of thought experiments regardless of their rational or empirical nature.
Comparing the Cognitive Success of Thought Experiments

So far, we have seen that platonic thought experiments are used to access to a kind of a priori knowledge gained through intuitions. In Norton’s empiricism, the cognitive reliability of thought experiments matches with the cognitive reliability of their argument form. According to an instrumentalist position – as Galileo’s use of falling bodies thought experiment – the main purpose of thought experiments is to achieve a specific cognitive outcome. In what follows, I will answer if it is possible to compare the cognitive success between two or more competing thought experiments.

To begin with, let’s think on the notion of scientific “confirmation” as a relation between three related scientific concepts i.e. evidence, thought experiments and accepted background knowledge. About the former concept, and according with a Popperian interpretation of the concept of “evidence”, less expected evidence has more “confirmation – corroboration – value” than more expected evidence. About the later concept, the “accepted background knowledge” is seen here as those beliefs allowed by a specific conceptual frame.

With this in mind, we can say that a qualitative theory of deductive confirmation might explain the idea of “confirmation” as the increasing plausibility of the evidence (E) provided by the cognitive outcome (CO) of a thought experiment (TE). Then, if we interpret the notion of “plausibility” as the conceptually sound basis of TE, we get the following confirmation rule:

**Rule 1**: E provided by CO confirms TE if and only if (iff) TE makes CO more plausible.

Note that the “conceptual sound basis” required here is determined by scientists’ no problematic accepted background knowledge. And of course, this is a lengthy debatable issue. Nevertheless, the rules of confirmation presented here are thought to be independent from the rational – and empirical – content of thought experiments. In other words, to what extent the evidence provided by the cognitive outcome of a thought experiment increases its conceptual sound basis, it is not a question that can be answered using these rules designed solely to evaluate the explanatory power of thought experiments in competition.

According to rule 1, if we compare two different and competing thought experiments with different CO, then we can infer the following confirmation rule:

**Rule 2**: CO confirms TE₁ more than CO* confirms TE₂, iff TE₁ increases the plausibility of CO more than TE₂ increases the plausibility of CO*.

If we compare the hypothetical case where a TE produces two possible incommensurable CO, then:

**Rule 3**: If TE makes CO as plausible as CO*, then CO confirms TE as much as CO* does.

If we compare the hypothetical case where CO equally confirms two different thought experiments, then:

**Rule 4**: If CO equally confirms TE₁ and TE₂, then CO confirms TE₁ more than TE₂, iff TE₁ is more plausible than TE₂.

Finally, we can build a general rule for thought experiment choice:
Rule 5: When TE_1 has so far proven to be more successful than TE_2, eliminate TE_2 in favor of TE_1 (at least for the time being).

Is my contention that these rules can serve to evaluate two or more competing thought experiments independently of the rationalist-empiricist debate in which the discussion of this type of experiment has been engaged in recent years.

Concluding remarks

If we grant that the “mind’s eye” capacity of scientists to “grasp” relevant laws of nature are equivalent to the notion of scientific “intuition”, Brown still should provide us with an epistemological justification to show that the use of scientific intuition is a reliable cognitive tool to grasp the metaphysics of natural laws he contends. In the case of Norton, he did not make the attempt to explain under which epistemological conditions we can assess if the knowledge presupposed in the premises of the argument form of thought experiments favored by him is true or false, which makes his position trivial. Finally, through Galileo’s falling bodies thought experiments, I show that there is another way to assess thought experiments beyond the rationalist and empiricist positions, i.e. the instrumentalist stance. This instrumentalist point of view allows us to develop some confirmation rules to compare the cognitive success between two or more competing thought experiments regardless of their rational or empirical nature.

References


Evaluating the Cognitive Success of Thought Experiments
Damián Islas Mondragón


