A Historical Analysis of Thought Experiments
Uma Análise Histórica dos Experimentos de Pensamento

Roberto Schmitz Nitsche
Mestrando em Filosofia pela Universidade Federal de Santa Maria (UFSM)
ronitsche@gmail.com
http://lattes.cnpq.br/0823474212738588

Abstract
This paper provides a historical study of thought experiments, from the birth of philosophy until Thomas Kuhn’s approach. It will be done through a bibliographic review. The first section explores the possible connection between the emergence of philosophy and thought experiments, considering whether cases presented by the first philosophers are thought experiments and whether they had any term to describe them. In the second section, ancient and medieval philosophy will be discussed, analyzing passages of Plato and Aristotle to investigate whether they made use of thought experiments. We will point out that all the definitions used by the discussed authors present problems. A possible solution is an earlier discussion in which one seeks to confront definitions of what thought experiments are. The third and last section examines some of the authors who preceded the contemporary discussion, including Ernst Mach, Pierre Duhem, Karl Popper, Alexandre Koyré and Thomas Kuhn drawing attention to some problems with their theories.

Keywords: Thought Experiments. History of Philosophy. Epistemology.

Resumo
Este artigo fornece um estudo histórico sobre experimentos de pensamento, desde o nascimento da filosofia até a abordagem de Thomas Kuhn, que será executado através de uma revisão bibliográfica. A primeira seção explora a possível conexão entre o surgimento da filosofia e dos experimentos de pensamento, considerando se os casos apresentados pelos primeiros filósofos são experimentos de pensamento e se eles possuíam algum termo para descrevê-los. Na segunda seção, discutiremos a filosofia antiga e medieval, analisando passagens de Platão e Aristóteles para investigar se eles fizeram uso de experimentos de pensamento. Apontaremos que todas as definições usadas pelos autores discutidos apresentam problemas. Uma solução possível é uma discussão anterior, na qual se busca confrontar definições sobre o que são experimentos de pensamento. A terceira e última seção examina alguns dos autores que precederam a discussão contemporânea, incluindo Ernst Mach, Pierre Duhem, Karl Popper, Alexandre Koyré e Thomas Kuhn, chamando a atenção para alguns problemas com suas teorias.

1. The Beginning of Greek Philosophy and its Connection with Thought Experiments

Thought experiments are not new to the modern or contemporary world. In addition to being able to find this type of tool being used in medieval philosophy, one can also find it in the early days of philosophy. Perhaps it is even possible that philosophy has made use of thought experiments since its birth. One can hardly prove this last point with total certainty, because of the little material that has managed to reach us after more than two millennia of philosophy.

Some works by Katerina Ierodiakonou shed light on thought experiments within ancient philosophy. Among her works are: “Ancient Thought Experiments: A First Approach” (2005), “Remarks on the History of an Ancient Thought Experiment” (2011) and “The Triple Life of Ancient Thought Experiments” (2018). In her most recent work, Ierodiakonou argues that our modern term “thought experiment” is not found in the ancient writings: “there is no ancient Greek term corresponding to what we nowadays refer to as a thought experiment, and presumably ancient philosophers did not have our modern notion of a thought experiment” (IERODIAKONOU, 2018, p. 31).

Even though there is no such ancient term that corresponds with the current one, Ierodiakonou argues that “there are ancient texts that refer to typical cases of thought experiments as ‘paradeigmata’, i.e., as ‘examples’” (2018, p. 32). In her interpretation she understands that:

The preposition “para” indicates that something is being put, placed, thrown beside something else for comparison or juxtaposition, while “deigma” means “sample,” “pattern,” “plan,” “model,” “sketch”; and so “paradeigma” must be understood, in its literal sense, to mean “the sample, pattern, plan, model, sketch that is placed beside something else for comparison or juxtaposition. (IERODIAKONOU, 2018, p. 32).

Ierodiakonou argues that the first recorded thought experiment, i.e. paradeigma, possibly the first in history, is attributed to Archytas, a philosopher who probably lived four centuries before our era. His thought experiment works on a model of reductio ad absurdum, that is, one accepts as a hypothesis exactly the opposite of what one wants to prove using the thought experiment, as this hypothesis leads one to an absurd conclusion, one ends up proving that the opposite is true.

The experiment developed by him asks us to imagine someone at the outer edge of the universe trying to reach his hand across it. Archytas offers us two possibilities: either the
person’s hand crosses or does not cross the edge of the universe. If a person succeeds in crossing his hand, then he can go further and try to do it again, and if he succeeds, then he can go even further, and he can go through this situation *ad infinitum*. On the other hand, if he cannot pass through his hand, there will be some kind of barrier, and that barrier must be made up of some matter. Therefore, if one assumes the hypothesis that the universe is finite, one arrives at an absurdity and concludes that the universe must be infinite. According to the philosopher, it would be absurd not to be able to cross our hand, but, on the other hand, if the person can cross his hand, then he would still not be at the edge of the universe.

Ierodiakonou does not provide an entirely clear argument for classifying the case of Archytas as a thought experiment. Is it enough to classify a case as a thought experiment when it is equivalent to a “paradigma”? Is it the case that all thought experiments are some kind of sample, pattern, plan, model, sketch that is placed beside something else for comparison or juxtaposition? It seems to be a very weak statement; thought experiments can deal with comparisons of more than two elements, or simply do not compare anything, where one just aims to conclude something. There does not seem necessary for all cases of thought experiments to be some kind of comparison or juxtaposition.

In addition to Ierodiakonou, Nicholas Rescher produced important works focused on the study of thought experiments through a historical approach. Rescher was one of the first to develop a work that approached the history of thought experiments through the beginning of Greek philosophy. The author defends the point that “the use of thought experimentation in philosophy is as old as the subject itself” (RESCHER, 1991, p. 32). The emergence of philosophy and the use of thought experiments are closely linked: “it is [the Greek nature-philosophers of Presocratic times] who invented thought-experimentation as a cognitive procedure and practiced it with great dedication and versatility” (RESCHER, 1991, p. 31).

Rescher uses the following definition of a thought experiment to analyze the thought experiments developed by the pre-Socratic philosophers:

> A “thought experiment” is an attempt to draw instruction from a process of hypothetical reasoning that proceeds by eliciting the consequences of an hypothesis which, for aught that one actually knows to the contrary, may well be false. It consists in a reasoning from a supposition that is not accepted as true – perhaps is even known to be false – but is assumed provisionally in the interests of making a point or resolving a conclusion. (RESCHER, 1991, p. 31).

Rescher’s definition is quite problematic. It can be very useful to talk about some specific class of thought experiments, but it certainly will not be fair if one takes it as a broad
definition, in a normative sense. It is not the case that all thought experiments are an attempt to draw instruction. What is the type of instruction that the cases presented by him draw (we will see them below)? The author further states that a thought experiment consists of reasoning from a supposition that is not accepted as true, perhaps is even known to be false. This clearly cannot be the case; if it were, one would not be able to produce any case in which one followed widely accepted and true assumptions.

The first pre-Socratic examined by Rescher is Thales of Miletus. Thales is said to be one of the likely first philosophers in history, having allegedly stated that “the earth floats on water (like a piece of wood or something of the sort)” and that “the ‘principle’ of all things is water” (KIRK, G.S.; RAVEN, J. E, 1957, p. 88). For Rescher, Thales follows the Method No. 1 (Explanatory Conjectures), in which it is sought to demonstrate that $P$ is the case, when it has not yet been established that $P$ or that $\sim P$. Thus, through thought experimentation it is assumed that $P$ is the case (which is not completely implausible), and with this assumption, it is explained that $Q$, where $Q$ is patently a truth which one could not readily explain otherwise. Therefore, it remains that $P$.

Following this method, Rescher converts the two passages mentioned above into thought experiments. To be proved: that the Earth floats in water like a log. One assumes that it is the case that the Earth floats in water like a log. This assumption explains why the Earth remains in place in nature (and does so as well as any alternative available to explain it). One is therefore justified in saying that the Earth floats in water like a log. Now “consider the case of the psyche, the principle of life, regarded as that which enables living things to be living, and which all living things thus have to have in common” (RESCHER, 1991, p. 33). To be proved: that psyche is made of water. Suppose that the psyche is made of water. Such an assumption explains why all seeds have moisture and need water for their growth. One is therefore justified in claiming that psyche is made of water.

Here one can clearly see a conflict of claims between Ierodiakonou and Rescher, she claims that the first recorded thought experiment was elaborated by Archytas, on the other side, he claims that the two cases of Thales of Miletus are thought experiments. Several potentialities of interpretation are possible. If one accepts that the two cases are legitimate examples of thought experiments, then Thales wins the day because he lived chronologically before Archytas. On the other hand, the case of Archytas seems to be closer to the most famous

---

1 An analysis of the discussion can be found at KIRK; RAVEN, 1957, p. 88.
examples of thought experiments, while the cases of Thales seem only to be arbitrary assumptions, where one can argue for the lack of the experimental element, which can, in turn, be found, to some extent, in the case of Archytas.

Rescher also commented Anaximander’s philosophy to prove that he also used thought experiments. For Anaximander, the Earth remains in place because of its balance. It is established in the center and is also related to the extremes, it must not be carried either upwards or downwards, not even sideways. In this way, it remains fixed by necessity, and cannot move in opposite directions simultaneously. The Earth is on high, supported by nothing, but remains on account at a similar distance from all things².

Rescher argues that Anaximander used what he calls Method No. 2 (negative demonstrative reasoning). The method seeks to demonstrate that P (where one still doesn’t know that P or ~P). One assumes through a thought experiment that ~P. From this assumption one deduces that Q, Q being a potentially false thesis. Therefore, one maintains that P. According to the author, “the reasoning thus takes the line that if the earth were not at the center, then it would eventually succumb to a tendency to move it further in one direction or another, and so would not have a stably fixed and firm position at all” (RESCHER, 1991, p. 34).

In his consideration of the Pythagoreans, the author argues in favor of an advance in the use of thought experiments: “the negatively probative mode of hypothetical reasoning came to be transmuted into a formal mathematical method of proof – the mode of demonstration that has come to be known as reduction ad absurdum argumentation” (RESCHER, 1991, p. 35). This new method (method No. 3 (reduction ad absurdum)) advances as follows: one tries to demonstrate that P. It is assumed through a thought experiment that ~P. This assumption leads to a complete contradiction. Therefore, it is established that P. Rescher declares that “the notorious proof of the incommensurability of the diagonal of a square with its sides was accomplished by just this device” (RESCHER, 1991, p. 35).

While exploring Xenophanes’ passages, Rescher argues that he also made use of thought experiments. Xenophanes argued that a process of mixing earth and sea was taking place; he proposed that the sea would be dissolving the earth. The opposite process, a solidification of the sea, would also have already occurred. As evidence, he points out that shells, seaweed and fish impressions can be found on rocks and mountains far from the sea. They were produced a long time ago, when everything was covered with mud, and, when drying, the mud was left

² For details of the discussion see KIRK, G. S.; RAVEN, J. E, 1957, p. 134.
with the impressions of these beings. Rescher points out that this would be another example of the use of method number one: “this passage clearly shows that Xenophanes sought to substantiate his doctrine of alternative phases of solidification and dissolution through the use of thought experiments by way of explanatory conjectures” (RESCHER, 1991, p. 36).

Xenophanes, on the other hand, would also have developed a new method of using thought experiments, called method No. 4 (Skeptical Thought Experimentation). The method is used in the following way: accepting that things are the way they are, one will be inclined to accept the truth of P. But, assuming through a thought experiment that our state is appropriately different, then one would not accept that P, but rather that P’, which is incompatible with P. Therefore, one is not truly guaranteed in the categorical acceptance of P. The example for the use of this method, according to Rescher, can be found in the passage in which Xenophanes asserts that if other animals, that are not human, had hands and could portray through drawings how they see their gods, each of the animals would embody their god with their shape, that is, horses would draw a god in the form of a horse, just as oxen would draw him in the form of an ox.

Rescher also analyzes another passage from Xenophanes in which he states that if God had not created honey, we humans would consider figs to be the sweetest thing. Things being as they are, honey is the sweetest thing in the world, the epitome of sweetness. But suppose that honey does not exist, so figs would be the sweetest thing in the world known to humans, figs would, therefore, be the epitome of sweetness. Therefore, one should not say that honey is the epitome of sweetness, it is just the sweetest thing one have known so far. The author argues that “Xenophanes relied on thought experiments to establish the relativity of human knowledge, a device that was later to prove a major armament in the arsenal of the skeptics” (RESCHER, 1991, p. 37).

The last pre-Socratic commented by Rescher is Heraclitus. He argues that, “of all the Presocratics, however, it was Heraclitus to whom thought experimentation came the most naturally. In his thought, the projection of ‘strange’ suppositions is a prominent precept of method” (RESCHER, 1991, p. 37). According to the author, Heraclitus developed two different methods for the use of thought experiments: method no. 5 (analogical thought experimentation) e method no. 6 (value dominance argumentation).

In the first method, one assumes that someone made X. Then, someone would say that someone is F (crazy or mad). However, when doing Y someone is also doing X concerning the
relevant considerations of F. Therefore, it should be said that when someone does Y he is also F. The author exemplifies this method with the following passage: “they vainly purify themselves by defiling themselves with blood, just as if one who had stepped into the mud were to wash with mud. Anyone who saw him doing this would deem him mad” (Frag. 5/129 & 130; Burnet, p. 145; Kirk & Raven, p. 211, in RESCHER, 1991, p. 38).

In the second method, one assumes, through a thought experiment, that X does not exist. Thus established, one could not form the concept of Y, with X and Y being related concepts such as hot and cold and cause and effect. Therefore, Y’s place in the total scheme of things cannot be less important or valuative than X. Rescher maintains that “Heraclitus uses this sort of reasoning repeatedly to argue for the mutual dependence of opposites” (RESCHER, 1991, p. 39), as seen in the following passage: “men would not have known the name of justice if these things [that people deem unjust] were not” (Frag. 23/60, Burnet, p. 137, in RESCHER, 1991, p. 39).

For Andrew D. Irvine, the examples provided by Rescher are not legitimate thought experiments. The author recognizes the historical importance that pre-Socratics have concerning thought experiments. However, they represent only a first step towards a more detailed future improvement. With this, Irvine does not want in any way to decrease the importance of the work they have done since they have introduced a curiosity about the natural world together with the importance of observation and the critical method. As much as the pre-Socratics made use of hypothetical reasoning, reductions to absurdity and explanatory conjectures, it does not follow “that the use of these forms of reasoning should, in itself, be considered sufficient to prove the existence of thought experiments” (IRVINE, 1991, p. 154). The use of these reasoning devices would represent only a first, but necessary, step towards what would later be labeled contemporaneously as a thought experiment.

The author accepts that when Thales argues that the Earth floats like a log to explain its stability, he is making use of hypothetical assumptions and also explanatory conjectures. However, Irvine wonders: “now, the question that must be asked is: do these two factors together mean that Thales’ conjecture should be understood to be a thought experiment?” (IRVINE, 1991, p. 154). Irvine claims it is not. It would be necessary for Thales to have more detailed observations about the world and a larger theoretical context for this example to be considered an experiment or even a thought experiment. In his words: “Thales’ conjecture remains simply a case of reasoning by analogy. As it stands, it just does not have enough that
is recognizably experimental about it to be considered a thought experiment” (IRVINE, 1991, p. 154):

What we are forced to conclude is that, although many of the types of reasoning used by Presocratics represent an important first step towards the development and use of thought experiments in science, this is all they represent, namely, a first step. Just as this degree of sophistication in experimentation in general was not found in the first Presocratic attempts at science, neither does it appear in their initial attempts at reasoning about the natural world. (IRVINE, 1991, p. 154).

Despite the definition of thought experiments proposed by Rescher, it seems that all the cases presented boil down to rules of logic (reductio ad absurdum, modus tollens, modus ponens, etc.), but, as pointed out by Irvine, is that enough? Both Ierodiakonou and Rescher start from problematic definitions of what thought experiments are, and, without having the appropriate and widely accepted definition, they venture to classify old cases as thought experiments. The whole discussion would be resolved, or at least it would be directed towards a solution that could be accepted by Ierodiakonou, Rescher and Irvine, if one has, first of all, a non-problematic definition of which cases count as thought experiments and which do not.

If one considers some cases, such as those pointed out by Rescher about Thales of Miletus, as being legitimate cases of thought experiments, then one can say that there is a strong link between the development of philosophical thinking with thought experiments. On the other hand, if one accepts, as Irvine says, that they represent no more than a first, but essential, step towards the development of thought experimentation, then one can say that the development of philosophy can be linked with the later appearance of thought experimentation, in whatever period one argues that this fact took place, for providing the main tools used in the execution of thought experiments.

2. Classical and Medieval Philosophy and Thought Experiments

Another interesting work on thought experiments about ancient philosophy was developed by Alexander Becker. The author analyzes some passages from the writings of Plato’s Republic to see if they can be classified as thought experiments. Becker seeks to elaborate on three necessary conditions on what can be categorized as a thought experiment. Here are the three restrictions:

1. Thought experiments are tools intended to establish a claim which in itself doesn’t belong to any fictional world. Their point is not just to build a fictional world. Instead,
their epistemic function is directed towards the actual world. If this constraint were dropped, any kind of fiction (myth, novel, or whatever) would be a thought experiment. 2. Thought experiments are not merely illustrations of something that could as well be established without them. If this requirement were lifted, any fictional example used in a philosophical text would count as a thought experiment. Of course, this constraint is context relative: the thought experiment has to be irreplaceable only within a given context. 3. Thought experiments must somehow be decidable. At least, they must invoke some means by which the acceptability of the claim they are to support can be decided, even if in fact the decision might not be feasible. Without this constraint, thought experiments would merely be more or less entertaining stories, but they would lack an epistemological payoff. (BECKER, 2018, p. 48).

Are Becker’s restrictions good ones? Why can’t one call a thought experiment a case within a fictional world that seeks to elaborate through thought experimentation a truth about that world? Can it not be the case that myths and novels also seek to establish truths about this world, in addition to truths about their fictional worlds? The second constraint conflicts with an approach that states that thought experiments are arguments (NORTON, 1991). It states that one can reconstruct all thought experiments in the form of arguments, no longer needing the thought experiment. Thus, thought experiments are merely illustrations of something that could as well be established without them.

Becker’s first analysis is about the Gyges ring, which he describes as follows:

Gyges was a shepherd in the service of the king of Lydia. One day, he found in a subterranean cave a ring which could make him invisible. By doing so, this ring allowed him to do and to get whatever he wanted without being discovered. What he wanted was, firstly, the king’s wife, and then, the king’s position. With the help of the ring, he obtained both easily. Now consider, Glaucon continues, two people, one just, one unjust, both given such a ring: how would they behave? In the end not so differently, says Glaucon. Even the just man could not resist the chance to take whatever he wants if he had no longer to fear any consequences. This is, so Glaucon’s conclusion, a sure sign or proof that nobody acts justly if not for the consequences of acting justly. (BECKER, 2018, p. 49).

Becker argues that the first requirement is met since “this story is intended to establish a thesis which is supposed to apply generally. It is the claim that justice is desirable only for its consequences” (BECKER, 2018, p. 49). The second requirement is also met, the story is not intended to merely illustrate a circumstance. The third is also met, because “Glaucon proposes that there are two possible motivations for acting justly: because justice is desirable in itself, and because it has desirable consequences” (BECKER, 2018, p. 49), that is, one can dispute between two possibilities.
The author also examines myths that can be found in Plato, but that cannot be considered as thought experiments because they do not meet the three requirements he elaborated. We will see below two of his analyses. The first is about the myth of Er:

The myth is narrated as the recounting of what Er’s soul experienced when it traveled through the otherworld. It traveled together with the souls of the deceased to the underworldly judgement place. There, it becomes witness of how some souls are sent downwards, others upwards, and it listens to the reports of those who return from these places. Then it follows those who have finished their time of punishment or reward and who are ready for another round of reincarnation; on this way, it also catches a glimpse of the structure which holds the world together. (BECKER, 2018, p. 50).

The first condition, according to Becker, is doubtfully satisfied, “on the one hand, the claim that justice has desirable consequences goes beyond the fictional sphere of the myth, on the other hand, the more specific claim that justice has desirable consequences after death” (2018, p. 50), thus, it is not possible to use verification and falsification methods. This also leaves doubts about condition two, which requires that one is not just seeing a mere description of an event or story. Condition three is not satisfied, one cannot dispute whether “we have the option to believe in what the story suggests, or not to believe in it; but this choice would be equal to accept or not to accept the use of thought experiments in the first place” (BECKER, 2018, p. 50).

Becker also explores the allegory of the cave. According to the author, the allegory of the cave “is about our way of education, learning and understanding if we want to pursue the road of philosophy” (BECKER, 2018, p. 50). In short, the allegory exemplifies the arduous process of acquiring knowledge and the thorny process of convincing other people to see the reality. In the allegory, people are in a cave where they can see only shadows, and this is the only reality they know. When one of the prisoners manages to escape, he goes through the process of being blinded by the sunlight. Upon returning to the cave and warning others, they vehemently scold him and decide to continue his comfortable life by seeing the shadows, instead of being blinded with the sunlight. Becker argues that the first and third conditions are satisfied, since the allegory is not only the construction of a fictional world, it seeks to bring something to the present world, and it is also decidable, since one has methods for the analysis of its veracity. However, condition two is not completely satisfied “due to the simple fact that similes are open to an interpretation which replaces the simile by an account which deals directly with the explanandum” (BECKER, 2018, p. 51).
Klaus Corcilius argues that like other authors of the time, Aristotle also had no conception of a thought experiment: “he does not discuss them in his works, and his writings do not show signs that he has identified a distinctive mode of pointing out something, philosophically or otherwise, that could plausibly be described in terms of what we call thought experiments” (CORCILIUS, 2018, p. 57). As much as one cannot find a concept or term that has an equivalence to our contemporary understanding of thought experiments “Aristotle undoubtedly makes abundant use of what we call thought experiments” (CORCILIUS, 2018, p. 59). The use of thought experiments in Aristotelian philosophy occurs in “cases where it would be impossible or exceedingly difficult to carry out physical experiments, but also, in cases where it would be relatively easy to perform the corresponding physical experiments” (CORCILIUS, 2018, p. 59). The author’s interpretation is that for Aristotle “thought experiments are a means to compensate for a lack of data in epistemically difficult terrain, by somehow generating data from imagined scenarios” (CORCILIUS, 2018, p. 60).

Within what Corcilius authorizes to be called a thought experiment are: “the stripping argument”, an experiment that seeks to know what substance is. The use of this thought experiment occurs as a way to refute other theses; “A second sun”, where Aristotle “critically confronts the Platonic doctrine of Ideas with difficulties” (CORCILIUS, 2018, p. 63); another thought experiment asks us to imagine that “all things are colors”. Aristotle uses this thought experiment against mathematical philosophers who “posit unity (“the one”) and numbers to be fundamental items of their ontologies and explain what things other than numbers are in terms of such metaphysical numbers” (CORCILIUS, 2018, p. 64), trying to “demolish that thesis by showing how number and unity are too thin as concepts to be able to account for what things other than numbers are” (CORCILIUS, 2018, p. 64). Yet another thought experiment tries to prove “why flesh is not the organ of touch” (CORCILIUS, 2018, p. 65). Corcilius also classifies as cases of thought experiments the cases in which Aristotle discusses “if white would be the only perceptible” (CORCILIUS, 2018, p. 66) and “refutation of the Atlas-theory” (CORCILIUS, 2018, p. 67).

Corcilius, therefore, presents a broader definition, within Aristotelian philosophy, of what is necessary for something to be defined as a thought experiment in relation to the requirements used by Becker for the same sort of evaluation in Platonic philosophy. Corcilius describes the Aristotle’ version of the cave allegory as follows:
Aristotle invites us to imagine people who, apart from living in underground dwellings without knowledge of the existence of outside world, live in circumstances similar to ours; they have everything they need, as well as some vague stories from hearsay about the existence of a divine power. Next he invites us to imagine that “the jaws of the earth opened” so as to expose the cave-dwellers to the sight of the stars and their regular motions. On that basis, he argues that the cave-dwellers, overwhelmed by the sight, would “most certainly” come to believe there to be gods, who, as only gods could, have created this marvelous universe. (CORCILIUS, 2018, p. 61).

As we have seen, Becker does not accept the cave allegory as an example of a thought experiment; Corcilius, on the other hand, accepts the cave allegory presented in Aristotle as being, in fact, a thought experiment, because “the thought experiment provides otherwise unavailable data, not by disclosing hitherto hidden facts, but by providing a new perspective on data that are all too obvious to everyone” (CORCILIUS, 2018, p. 61).

About the medieval period, we still do not have a significant amount of work on the use of thought experiments. Among those available is an article by Jon McGinnis on thought experiments in the medieval Islamic world, within which he mainly considers Avicenna’s philosophy. McGinnis states that there was no Arabic term for thought experiments, however, they would have been widely used by philosophers at the time:

There is no (medieval) Arabic term or phrase for “thought experiment.” Be that as it may, medieval philosophers and scientists working in Arabic both concretely employed thought experiments in their philosophies and discussed their merits and demerits abstractly. Indeed, it would seem that thought experiments truly captured the imagination of medieval thinkers in the Muslim world, who left behind a significant body of examples and analyses of such experiments. (MCGINNIS, 2018, p. 77).

Another paper, “Mediaeval Thought-experiments: The Metamethodology Of Mediaeval Science”, on the topic was developed by Peter King. The author points out that, in general, there was little concern among the philosophers of the High Middle Ages regarding whether the cases used in their philosophical theories could or could not be concretely realized. The author argues that, in general, “the only unifying mark all these cases have, I believe, is that they are thought-experiments” (KING, 1991, p. 49). Therefore, “thought-experiment is the methodology of medieval science” (KING, 1991, p. 49).

King also seeks to demonstrate that medieval disputes called obligatio are related to thought experiments. The author characterizes it as follows:

A typical obligatio has the formal characteristics of a debate or mediaeval dispute: There are two parties, an opponent and a respondent. The opponent begins by laying down some claim, a proposition such as “Socrates is running”. His action is called position, positing, and what he points is the casus, the case, also known as the positum. The respondent admits the case – or, if he does not, there is no dispute – and then the
opponent proceeds to put forward (to propose) other propositions. To each proposition, the respondent either concedes, denies, or “doubts” its truth; he is “obliged” to give certain responses in accord with the rules; hence the name obligations. The point of the exercise, if there is an identifiable point, is to trap the respondent in a contradiction; an obligational dispute explores “what happens” given the positum. (KING, 1991, p. 51).

About the multiple rules that could be stipulated for this debate, there are at least three that, according to the author, “constitute a theory of thought-experiments” (KING, 1991, p. 52). The first point indicates that there may be the use of positum that does not need to be possible: “this corresponds, in thought-experiments, to reasoning per impossible.” (KING, 1991, p. 52). The second point states that obligatio uses counterfactual reasoning in the sense that “[it] takes contingent facts in the actual world as relevant to determining what would happen under a given condition” (KING, 1991, p. 52). The third point is about the inclusion of semantic content: “several authors allow for distinct species of position, in cases in which the positum includes explicit semantic content. Theses varieties were useful in investigating logical and linguistic issues” (KING, 1991, p. 52).

King’s description seems to be quite fruitful and accurate concerning what most people understand by thought experiment. They allow us to postulate sentences that one may not actually accept and use them to perform reasoning per impossible, and with that, one obviously ended up using counterfactual reasoning. However, this approach works only if one takes it as descriptive and not normative, as a thought experiment doesn’t necessarily need to use counterfactual or per impossible reasoning.

King asserts that thought experiments in the medieval era have their virtues “appropriate for investigating a priori truths and uncovering conceptual incoherencies and inadequacies” (KING, 1991, p. 56). Thought experiments were used in very rudimentary ways during the Middle Ages, often to support “theories which have no check or control, no way to test their correctness or incorrectness” (KING, 1991, p. 56), if compared with the way one uses them in contemporary times.

3. The (Re)Discoveries of the Importance of Thought Experiments
According to the entry on thought experiments developed by Brown and Fehige in the Stanford Encyclopedia of Philosophy, there were at least four stages of (re)discovering the importance of using thought experiments:

Accordingly, the modern history of the philosophical investigation into thought experiments can be divided into four stages: in the 18th and 19th century the awareness of the importance of thought experiments in philosophy and science emerges. In addition to Lichtenberg, special mention should be made of Novalis and Hans-Christian Ørsted. The topic reemerges in a more systematic manner at the beginning of the 20th century with little relation to the attempts made at the first stage. The stakeholders of the second stage were Pierre Duhem, Mach, and Alexius Meinong. A third stage, probably due to the rediscovery of the importance of scientific practice for a proper understanding of science, followed in the first part of the second half of the 20th century. Again, the contributions of this stage bear little relation to the two previous stages. While the third period has seen a number of noteworthy contributions, the protagonists of this period were Alexandre Koyré, Thomas S. Kuhn and Karl Popper. The ongoing philosophical exploration of thought experiments began in the 1980s, and marks the fourth stage. Arguably, it has been the most prolific one of all four stages. With some very important sign-postings the ongoing discussion took off after 1991. James Robert Brown and John D. Norton have carried on a debate that others find useful, especially to contrast with their own alternative accounts.

It is in the first stage that a term is finally developed to refer to thought experiments. Some authors claim that the term first appeared in the writings of Ernst Mach (Gedankenexperiment), however, it was Hans Christian Ørsted “who coined the term within the context of German Naturphilosophie. Also, a term for experiment with thoughts (mit Gedanken experimentieren) [was] found, a few years ago, in a 1793 entry to German polymath Georg Christoph Lichtenberg’s ‘Common place book’” (ASPASIA, S. M.; KYRIAKOS, A. M.; HAIDO, K., 2006, p. 63).

Ernst Mach was one of the rediscoverers of the essentiality of thought experiments. In one of the chapters of his book Erkenntnis und Irrtum he dedicates a chapter to talk about thought experiments and their significant connection with science: “experiments guided by thought lie at the basis of science and consciously aim at widening experience” (MACH, 1976, p 135). However, he also claims that they are valuable not only for the sciences: “we can hardly doubt that thought experiments are important not only in physics but in every field” (MACH, 1976, p. 143). In the same way, thought experiments are also central for the development of our reasoning ability, one can develop a good aptitude in abstract reasoning by making frequent use of thought experimentation: “experimenting in thought is important not only for the

---

3 The entry used in the following quotes refers to a version before to the substantive revision made to the entry on September 26, 2019. Unfortunately, the quotes used in this paper are not found in the new version of the entry.
professional enquirer, but also for mental development as such” (MACH, 1976, p. 143). Mach even claims that this was one of the most key factors for the construction of science as one knows it today: “the close conjunction of thought with experience has built modern natural science” (MACH, 1976, p. 146).

For Mach, thought experiments are imagined conditions that connect our expectations and assume certain consequences. For him, thought experiments can be of two types: those that are fanciful and combine certain conditions that will never occur in reality and those whose ideas are good representations of the facts, they come very close to reality. The first type is performed by planners, aerial castle builders, novelists and builders of social and technological utopias; while the second type is performed by stubborn traders, serious inventors and inquirers4.

In this approach, thought experiments are more easily manipulated when compared to physical experiments and have a lower cost, in several ways, to be performed. According to Mach’s understanding, therefore, it is natural that thought experiments often precede and set the stage for the execution of physical experiments. However, thought experiments are not necessary conditions for carrying out physical experiments. Every experimenter needs to properly arrange the whole scheme of the experiment in his mind before translating it into facts: “Galileo must see the experimental arrangement for investigating free fall well represented in his phantasy before he can realize it” (MACH, 1976, p. 137).

According to Mach, the functioning of thought experiments occurs through memory that helps us to identify details of the real world that one failed to give due attention to before starting the thought experiment, in this way one can reach new conclusions and knowledge through its use:

For we can find in memory details that we failed to notice when directly observing the facts. Just as in memory we may discover a trait that suddenly reveals a man’s character hitherto misread, so memory offers new and so far unnoticed features of physical facts and helps us to new discoveries. (MACH, 1976, p. 132).

But this is really false, if it wants to be a complete theory. There are many cases of thought experiments that do not use any facts that one has previously experienced. Thought

4 Brendel argues that “this very general conception of thought experiments does not adequately reflect the fact that a thought experiment is indeed a certain kind of experiment. Although it is an imaginary investigation that need not or cannot be executed in the real physical world, it is nevertheless subject to certain theoretical requirements that it shares with real experiments. For example, a ‘thought experimenter’ also studies the functional dependency of variables by planned and controlled data change. Furthermore, in a manner similar to real experiments, every thought experiment depends on some background assumptions or background theories”. (2004, p. 91).
experiments are known for using counterfactual reasoning and possible worlds. One imagines situations where laws of physics are suspended or where reality is slightly or completely changed. One certainly never experienced any of these cases, and therefore one has no memories about them to turn to when executing thought experiments.

In a certain sense, in Mach’s view, thought experiments play a lesser role than physical experiments. Thought experiments “can be so definite and decisive that the author rightly or wrongly feels able to dispense with any further tests by physical experiment (MACH, 1976, p. 137). On the other hand, one still needs the physical experiment to achieve greater confidence in the result: “the less certain their outcome, the more strongly thought experiments urge the enquirer to physical experiment as a natural sequel that has to complete and to determine the result” (MACH, 1976, pp. 137-138). In this way “it is small wonder that thought experiment often precedes and prepares physical experiments” (MACH, 1976, p. 136). The physical experimenter must readily have an entire scenario available in his mind before carrying out the physical experiment: “thought experiment is in any case a necessary precondition for physical experiment. Every experimenter and inventor must have the planned arrangement in his head before translating it into fact” (MACH, 1976, p. 136).

Thought is at a lower level concerning experience, “the thought usually contains less than experience” (MACH, 1976, p. 137), and in thinking, one usually finds only schemes about reality that may contain premeditated additions; nevertheless, it is in thought that one can, through the use of the memory of experiences, reorganize what one has already experienced and then “[one will be able] to learn how accurately experiences are represented by thoughts and how far the latter agree with each other” (MACH, 1976, p. 137). It is through this recombination in thinking that one can improve the scope of our ideas and make the thought experiment more reliable: the basic method of thought experiments being variation, that is, “by varying the conditions, the scope of ideas (expectations) tied to them is extended: by modifying and specializing the conditions we modify and specialize the ideas, making them more determinate, and the two processes alternate” (MACH, 1976, p. 139).

As noted by Brendel, Mach’s definition of thought experiments is still very modest concerning all the philosophical arguments that will be formed in later years. Brendel argues that “Mach uses the term in a very wide sense. According to him, thought experiments can be almost all kinds of ‘thought experiences’, like dreaming, hallucinating, writing novels or imagining utopia” (BRENDEL, 2004, p. 90-91).
Ernst Mach’s “Erkenntnis und Irrtum” was published 1905; Pierre Duhem’s book “La Théorie Physique. Son Objet, sa Structure” appears the following year, in which there is a strong skepticism about the use of thought experiments. In one of the chapters entitled “Consequences Relative to the Teaching of Physics”, Duhem criticizes the idea that “it is generally accepted that each hypothesis of physics may be separated from the group and subjected in isolation to experimental test” (DUHEM, 1991, p. 200). The author argues that people would like the teacher to organize the hypotheses in a certain way, in order, exposing experimental checks and, later, declaring them accepted; thus, passing from hypothesis to hypothesis until all physics was constituted.

Duhem argues that this type of fictional experiment leads us into a vicious circle:

To invoke such a fictitious experiment is to offer an experiment to be done for an experiment done; this is justifying a principle not by means of facts observed but by means of facts whose existence is predicted, and this prediction has no other foundation than the belief in the principle supported by the alleged experiment. Such a method of demonstration implicates him who trusts it in a vicious circle. (DUHEM, 1991, p. 202).

These fictional experiments do not bring us any precise results: “the very indecisive and rough results it would produce could undoubtedly be put into agreement with the proposition claimed to be warranted” (DUHEM, 1991, p. 202). Another problematic feature pointed out by Duhem says that these experiments are often not able to become real experiments, as they make use of elements that one cannot find in nature. Duhem maintains that “what we have said suffices to warrant the following conclusion: the teaching of physics by the purely inductive method such as Newton defined it is a chimera” (DUHEM, 1991, p. 203). He argues that the need for these experiments to be carried out physically “[is] not the base of theory, but its crown” (DUHEM, 1991, p. 204).

As pointed out by Aspasia, Kyriakos and Haido, the skeptical trend towards thought experiments remained active through the Vienna Circle, where its members attributed “only a very minor role to thought experiments” (ASPASIA, S. M.; KYRIAKOS, A. M.; HAIDO, K., 2006, p. 65):

Once the project of the logical empiricists is identified, their limited interest in thought experimentation is no longer surprising, because, as Reichenbach put it, the actual thinking process is, for them, not a legitimate subject of epistemology. The logical empiricists are not interested in ‘actual thinking’, imagining and cognition and their role in scientific discovery. With the chosen outlook of rational reconstruction of science in terms of logic, there did not appear to be any space for the consideration of
We know that the Vienna Circle was probably active between 1922 and 1936, following Duhem’s death in 1916. Albert Einstein, who lived between 1879 and 1955, that is, during a skeptical period regarding the use of thought experiments in science, made plentiful use of them in his physical theories. The third rediscovery of the importance of thought experiments begins to be formulated at this period. The main proponents of this new wave of studies of thought experiments are Popper, Koiré and Kuhn.

Karl Popper, in 1934, published “Logik Der Forschung”, in which in one of his appendices, “On the Use and Misuse of Imaginary Experiments, Especially in Quantum Theory”, he briefly passes on the subject of thought experiments, which are for him merely called imaginary experiments. The author argues that there are ways to make good and bad use of them; Popper points out that Galileo’s falling bodies experiment “[is] a perfect model for the best use of imaginary experiments. It is the critical use. I do not wish to suggest, however, that there is no other way of using them. There is, especially, a heuristic use which is very valuable. But there are fewer valuable uses also” (POPPER, 2005, p. 465). However, even though there are fewer valuable ways of using them, Popper does not want to create “the impression that [he is] doubting the fruitfulness of imaginary experiments” (POPPER, 2005, p. 464).

Critical thought experiments are developed with the aim of criticizing a theory in an attempt to show that some possibilities may have been overlooked. However, even in this type care should be taken with the use of idealizations in the defense of the theory in question: “it is important not to introduce any idealizations or other special assumptions unless they are favorable to an opponent, or unless any opponent who uses the imaginary experiment in question would have to accept them” (POPPER, 2005, p. 466). This restriction is pointed out by Popper to avoid apologetic use: “the main purpose of this note is to issue a warning against what may be called the apologetic use of imaginary experiments” (POPPER, 2005, p. 466).

According to Aspasia et al., Popper’s view fits in with his general idea that experiments can only refute hypotheses, but not confirm them:

A critical thought experiment is intended to show that some view, way of thinking, or background assumption is wrong. An heuristic thought experiment is intended to show what is the correct view or way of thinking. It’s worth noticing that this general proposal (that thought experiments can refute, but cannot confirm) is an exact analogue in the area of thought experimentation to the Popperian well-known view of ordinary experimentation and its relation to scientific hypotheses (i.e., that
experiments can refute hypotheses, but they cannot verify or confirm them). (ASPASIA et al., 2006, p. 66).

In 1968 Koyré publishes “Metaphysics & Measurement”, having a chapter dedicated to the exam of Galileo’s thought experiment on falling bodies. The author argues in favor of the historical importance that thought experiments had in scientific development. This importance is generated by the complicated process of carrying out physical experiments, which often involve the use of complicated and very expensive equipment. Another problem with real experiments is their lack of precision which leads us to a certain element of doubt about their results. Koyré points out that “it is impossible in practice to produce a plane surface which is truly plane; or to make a spherical surface which is so in reality. Perfectly rigid bodies do not exist; nor can perfectly elastic bodies; and it is not possible to make an absolutely correct measurement” (KOYRÉ, 1992, p. 45).

These inaccuracies lead us to have a gap “between empirical fact and theoretical concept” (KOYRÉ, 1992, p. 45) that cannot be filled. The author argues that this is the moment when one can make use of imagination, and, thus, one can create experiments in thought that deal with the ideal or even the impossible, that is, one can make use of theoretically perfect concepts that are not available for real experiments: “thus, it rolls perfect spheres on perfectly smooth, perfectly hard planes; it hangs weights from perfectly rigid weightless levers; it causes light to be emitted from point sources; it sends bodies to move eternally in infinite space”(KOYRÉ, 1992, p. 45). However, as much as one can achieve precision with this type of help, “it is not prevented from being wrong sometimes” (KOYRÉ, 1992, p. 46).

The best development on thought experiments that one can find in this third generation is contained in the book published in 1977 by Kuhn entitled “The Essential Tension”. In it, probably for the first time, there is a concern with the “imagined situation”, which would later be labeled by the word scenario; the author puts the concern of the experimenter to be familiar with the information put on the scene, and this information is prior to performing the thought experiment, taken from the physical world. Kuhn argues that during the execution of the thought experiment, no new information is inserted into it. Their main function is to make the scientist learn something about his own mental apparatus, that is, a correction occurs in the scientist’s conceptual system. In Kuhn’s words:

If a thought experiment is to be effective, it must, as we have already seen, present a normal situation, that is, a situation which the man who analyzes the experiment feels well equipped by prior experience to handle. Nothing about the imagined situation
may be entirely unfamiliar or strange. Therefore, if the experiment depends, as it must, upon prior experience of nature, that experience must have been generally familiar before the experiment was undertaken. This aspect of the thought-experimental situation has seemed to dictate one of the conclusions that I have so far consistently drawn. Because it embodies no new information about the world, a thought experiment can teach nothing that was not known before. (KUHN, 1977, p. 252).

The thought experiment, therefore, creates a conflict or a contradiction in the way the scientist thinks. In this sense, the thought experiment acts as a substitute of confused and contradictory concepts for new ones that fit the world. There is no defect inherent in the concepts, but in the way they can be applied to the structure of the world. Thus, “from thought experiments most people learn about their concepts and the world together” (KUHN, 1977, p. 253). This process is called by Kuhn reconceptualization: “this process I have elsewhere labeled Scientific Revolution” (KUHN, 1977, p. 263). All the necessary information for such a revolution was already contained in the scientist’s conscience, which then goes through a moment of crisis, as there is a contradiction or conflict of concepts, which brings this information to the attention of the scientist, who goes through a revolutionary process of reconceptualization. From that point, the scientist can see this same information in a new way. This is the central process of thought experiments: “a crisis induced by the failure of expectation and followed by revolution is at the heart of the thought-experimental situations” (KUHN, 1977, p. 263).

Although Kuhn accepts the existence of thought experiments that cannot be carried out physically, the conflict of concepts that they bring to the scientist must be those that nature itself could present. This conflict must be familiar to the scientist, he must have had some sort of confrontation with them before the thought experiment; if that requirement is not met, the scientist will not be ready to learn anything through them.

Final Considerations

5 Brown does not agree with this aspect of the theory developed by Kuhn because “there are several thought experiments which have nothing to do with detecting problems in an old theory (e.g., Stevin’s inclined plane or Newton’s bucket)” (BROWN, 2011, p. 111-112). Brown further states that “incommensurability problems do not seem to be present in the Galileo case. There has been no change of meaning in the terms ‘light’, ‘heavy’, and ‘faster’” (BROWN, 2011, p. 112).

6 Lukáš Bielik (2004, p. 4) argues that Kuhn’s approach fails to account for cases such as the Galileo’s thought experiment on the speed of falling bodies. The author argues that “this thought experiment had been far from being just a memory device for someone working in an ‘aristotelian paradigm’. For Galileo, it didn’t represent an anomaly. Rather the scenario presented in this experiment became a lucid evidence for taking seriously the idea expressed later as Newton’s first law of motion”.
Ierodiakonou offers us a beautiful work in looking for an ancient concept that can be compared with experimentation in thought, but, as we pointed out earlier, this is not enough to guarantee that these cases are in fact thought experiments. Problems are also faced by the definitions developed by Rescher, Becker and King, however fruitful they may seem. All of these authors end up facing a common problem: there is no good definition, descriptive or normative, about what thought experiments are or should be. Arbitrary definitions are taken and old cases are judged with them.

The problem may be resolved, or at least be directed towards a resolution, if the discussion turns to define what thought experiments are and what are not in the first place, before the discussion about which old cases can or cannot be classified as thought experiments and which is the oldest that has been recorded. It must not just be the case to create an arbitrary description or standardization that has not gone through some discussion process with other descriptions or standards, something that none of the authors presented here has done.

We have seen that seeking to base thought experiments only on memory, as Mach sought to do, does not work, as not all cases of thought experiments deal with something from memory. Many cases put one in scenarios that one has never been and one does not have any type of experience. After Mach, we explored the skeptical phase of Duhem and the Vienna Circle. A time when some change begins to occur with Pooper accepting thought experiments to some extent and later with Koiré pointing out their importance, where one can have ideal situations that physical experiments often fail to offer. We later examined Kuhn’s approach, which, despite being one of the best developed so far, still presents problems. Not all thought experiments involve refuting theories, and they are not always an anomaly in the scientist’s conceptual framework.

The contemporary discussion started around 1986 with the publication of the paper “Thought Experiments Since the Scientific Revolution” by James Robert Brown. Two other elements that guided the discussion were the conference held at the Center of Philosophy at the University of Pittsburgh in 1986, and the volume “Thought Experiments in Science and Philosophy” in 1991, composed of articles presented at the conference and commentators. In this volume is found the first article by John Norton, who would become the main critic of Brown’s approach. However, the discussions that originated from these events and that became the center of the current debate are too extensive to be discussed here and certainly deserve a paper in itself.
References


Recebido: 18-03-2020
Aceito: 22-12-2020