

# The Importance of Thought Experiments in Understanding Scientific Theories

**Juan Carlos Vélez Rengifo**

Departamento de Ciencias Del Comportamiento, Universidad Católica Lumen Gentium, Cali, Colombia

**Email address:**

[jvelez@unicatolica.edu.co](mailto:jvelez@unicatolica.edu.co)

**To cite this article:**

Juan Carlos Vélez Rengifo. The Importance of Thought Experiments in Understanding Scientific Theories. *Science Journal of Education*. Vol. 9, No. 2, 2021, pp. 32-39. doi: 10.11648/j.sjedu.20210902.12

**Received:** February 8, 2021; **Accepted:** April 8, 2021; **Published:** April 23, 2021

---

**Abstract:** Against the informational and propositional character of knowledge, the idea that understanding is not a form of knowledge will be developed, since it is related to veritism and reliability, while understanding (comprehension) has to do with explanatory, objectual, technical relations, as coherentism had suggested for the problem of justification. The turn in epistemology has to do with the problem of the value of knowledge, now in terms of understanding, since this gives order and systematicity to thought. Then it will be argued that the fundamental feature of science is representational capacity, and the means to represent are not true beliefs but modeling, that is, functional theoretical models as idealization of the most relevant characteristics, leaving aside those that are not, depending on the objectives. Finally, the thesis is defended that mental experiments do not generate knowledge but rather broaden scientific understanding, understanding this as the most important cognitive achievement. In other words, mental experiments are a narrative way of generating mobile models of scientific theories. Mental experiments should be understood as actions of thought that allow to deepen a theory, to refute a hypothesis, but mainly have didactic functions of explanatory nature. Mental experiments contribute to understanding and rely heavily on imagination and narrative. Mental experiments are a very useful cognitive tool since, not being true, they allow fictitious approximations that can be refined in terms of their adequacy to the theoretical models.

**Keywords:** Scientific Theories, Understanding, Thought Experiments, Scientific Models, Cognition

---

## 1. Classical Approach to the Problem of Truth

In his classic book "Believe, Know, Understand", Luis Villoro affirms that according to the semantic concept of truth, this consists of a relationship between two terms: sentences and facts; but in this conception the subjects do not appear nor are they necessary. The truth of a P sentence might be true even when people ignore it. However, knowing being an internal state of someone, when speaking of a known truth the subject must appear. Someone must judge the content of a proposition in relation to the facts and describe them as an asserted sentence. That is, what the subject does with a sentence is an assertion, that is, to judge that a fact exists or not.

If a sentence P is true, what P refers to exists independently of any subject, but I cannot know that something is true regardless of the way I apprehend the truth.

Objectively sufficient reasons or objective justification is what ensures for every subject that the object of his belief not only exists for him, but is valid for other people regardless of his judgment. Objectively sufficient reasons are those that put the criterion of truth and guarantee that beliefs will not fail.

For Villoro [15], objective justification implies the notion of truth that is correlated to reality and this is necessary to understand what objectivity is. Thus, objectivity comes to light when any subject can assert that the object of his belief has a real existence. But since objectivity consists in the coincidence of judgments of an epistemic community, the explanation of this coincidence has to do directly with the real existence that is independent of the subjects. Truth as correspondence, Tarski's thesis, results in this approach in a relationship between a judgment and reality, thus becoming the only meaningful rational explanation of objectivity.

Knowledge, focused as knowing, in that sense has a community dimension; This is the characteristic of scientific

knowledge, in this way, the definition of science that emerges from here is the following: "Science consists of a set of knowledge shared by a determined epistemic community: theories, statements that put them in relation to an und domain of objects, intersubjectively verifiable observation statements; All of this constitutes a body of propositions based on objectively sufficient reasons "(Villoro, [15]). We are going to oppose this conception of science, as well as the notion of knowledge that is concealed here. Let's look at where the problem of knowledge arises and that it ends up giving way to understanding.

## 2. Knowledge Is Not an Exclusive Achievement of Science

The problem of knowledge has currently been transformed into the problem of epistemic virtues, gradually abandoning that of how to justify beliefs to obtain true declarative sentences, and this has had an interesting consequence: evaluating the value of knowledge has led to thematization of value of understanding. I clarify at the outset that understanding is a interchangeable term with comprehension, I will use them interchangeably even if the latter is more common in our language. The problem begins in a dialogue by Plato, the Meno in which the value of true opinion is discussed against knowledge:

Socrates: Look, suppose someone knows the way to Larisa (or anywhere) and is on their way to that place while showing others how to get there; he will obviously be a good guide when it comes to that path.

Meno: Yes, of course.

Socrates: And what would happen to someone who only has a mere opinion of how to get there - that is, a correct opinion of how to get there - but has never been to that place, and does not know how to get there; Won't you also be able to show others that way?

Meno: Yes, of course.

Socrates: [...] Will he not be as good a guide with his true belief as he who possesses knowledge?

[Meno agrees]

Socrates: So, in other words, doesn't a correct opinion accomplish the same good that knowledge does? (cited by Fricker, [4])

The question that opens the entire discussion of contemporary epistemology follows:

Meno: Except for one point, Socrates: if we have knowledge, we will always achieve our goals; But if we only have a correct opinion, sometimes we will achieve our objectives, but sometimes not.

Socrates: What makes you say that, Meno? If one always has a correct opinion, will we not always achieve our objectives as long as we have that opinion?

Meno: Yes, good observation, Socrates... it seems that this is true; which leaves me wondering: if this is the case, then why do we value knowledge more than mere true opinion, and why are they treated as two different things? (cited by

Fricker, [4]).

Indeed, the question that would occupy a large part of epistemology and whose answers led to the development of theories of justification as a means of differentiating knowledge from true beliefs remains open. Veritism or verificationism, coherentism or reliabilism are an example of this. Is knowledge better than true beliefs? Tradition responds with a yes in unison, since we can have true beliefs of fluke or epistemic luck.

But recent epistemology has put the emphasis on the problem of the value of knowledge and very interesting questions have arisen: but what happens when we think about trivial or immoral knowledge? Counting grains of sand on a beach or the knowledge of how to stir up hatred at a presidential candidate, or how to instigate genocide are examples that knowledge has no more value than true beliefs.

Reliabilism in light of this approach suggests that the reliability of the source of beliefs is what adds value to knowledge, that is, it was argued that knowledge is a true belief which is the product of "reliable faculties or processes of belief formation (Zagzebski, [16]). However, the counterexamples were not long in coming: a fallible espresso machine is good because espresso is a good thing, consequently, the value of espresso (product) makes the reliability of the machine a good thing, but the fact that the machine is reliable does not add anything to the value of the coffee, that is, if the espresso tastes good, there is no additional difference if it comes from an unreliable machine; and analogously: if the belief is true, it does not make a difference if it comes from an unreliable belief-producing source (Zagzebski, [16]). Hence, the truth related to a source of truth does not explain the value of knowledge, and this because knowledge cannot be conceived as the external product of a valuable cause. To say it of another form: no value can be attributed to the effect by virtue of its causes. If the causes of beliefs function properly they have a value, but it does not follow from this that it is automatically a value in the very state of knowledge. In conclusion, it is not possible to identify knowledge with true beliefs originated by valuable or reliable causes: this is the machine-product model of beliefs. But then, where do true beliefs get their value? From justification?

Knowledge has traditionally been considered as justified true belief, and this excludes the problem of value, since the belief is endowed with the property of being justified. Investigating the problem of the value of knowledge takes us far from our interests. Let us now consider the difference between knowledge and understanding.

Epistemology has extended the problem of the justification of knowledge in response to the challenges of skepticism, by defining knowledge or justification in terms of virtues; But the problem with this approach is that it leaves out other valuable aspects of cognition such as wisdom or understanding. Understanding is not a kind of deeper knowledge, it is a cognitive achievement other than knowledge.

But particularly the type of knowledge that interests us is

scientific knowledge. Science, is generally accepted, is organized knowledge, which implies that it is made up of true beliefs, but its main product is not knowledge (Elgin, [2]). Science is one of the main cognitive achievements. What then are the cognitive achievements of science?

Can you have knowledge without understanding? Kvanvig [6] for example, points out that it does; for example, you can meet the dean of the faculty without understanding him; Likewise, it is possible to contrast the knowledge of chains of facts about an object of study without understanding them. For Kvanvig knowledge consists of loose fragments of information in the form of individual propositions, only when they come together they become understanding. This idea is also shared by Elgin when he affirms that the objects of scientific knowledge are individual facts that are in turn expressed in true propositions canted in true declarative sentences. This is the thesis of veritism, namely the idea that a subject's knowledge consists of discrete grains of information backed up separately.

Like Kvanvig, Elgin argues that science is holistic, therefore it does not consist of an aggregate of separate and independently supported assertions. Science is an "integrated and systematically organized explanation of a domain. We call that kind of explanation a theory" (Elgin, [2]). This point of view had already been defended by Quine when he stated that "theories face the court of sensory experience not individually, but only as a collective body" (Quine, quoted by Elgin, [2]). Thus scientific theories do not possess the trait of granularity, as well as knowledge.

However, accepting holism and affirming that the truth applies only to conjunctions of propositions that as a whole constitute the theory does not solve the problem of the cognitive value of science. The problem detected by Elgin is that in a conjunction of two propositions, if there are anomalies in the conjunction (for example, one of these is false and the other true), it makes the cognitive results not knowledge. But if these anomalies are insignificant or are not detected at the beginning, or are derived from the misunderstanding of a certain phenomenon, to discard them outright would be to lose valuable information for the theory. In Elgin's words: "The hopelessness of selectively eliminating falsehoods and false implications of a theory weakens the plausibility of claiming that scientific knowledge is what subsists when the falsehoods of a theory have been suppressed" [2]. In conclusion, if what we want is to clarify the cognitive contribution of science, knowledge does not represent the epistemic magnitude for it. So since good science does not satisfy the requirements of knowledge, what does it offer?

When we evaluate a scientific theory we should not ask if it expresses knowledge, but rather if it communicates understanding of the phenomenon, or if it is a good way of thinking or representing a phenomenon, object, topic or area of study, given that what we want is to understand. What science offers is a unified, integrated and evidence-based understanding of certain phenomena.

### 3. Understanding as a Value of Science

The term understand is related to knowledge. Con-having knowledge in relation to context implies that understanding and knowledge has to do with constructive effort. Understanding consists of a mental operation in which it is not only enough to have knowledge but it must be maintained, conserved, repeated and applied by the subject (Tuffanelli, [12]).

Jonathan Kvanvig in a chapter of his book "The Value of Knowledge and the Pursuit of Understanding", 2004 entitled "Knowledge and understanding", begins by approaching the term of understanding or understanding through grammatical forms, namely,

The use of the noun form in phrases like:

"He lacks an understanding of the baroque to make a good interpretation of Bach"

"My understanding of Heidegger is superficial"

In the adjectival form they are:

"He has an understanding attitude"

And in the verb form they are:

"Understand quantum theory"

"Understand that Petro could have been president"

"We understand that Duque is sinking public education in Colombia"

Of all these grammatical uses that epistemologists focus on are those that have to do with the study of cognitive achievements and successes, but more specifically those that Kvanvig will deal with are "when it is affirmed that there is an understanding of an object, such as a subject of study, and when the use of the term supposes understanding that something is the case [6], as in the case " he understands quantum theory ", that is, what does this understanding consist of that is different from knowing.

Understanding that something is the case in comprehend why, when, where and what; in each case there is a special kind of comprehension that is related to some truth that explains it correctly; the uses of the explanation are then what is relevant above all in the type of objectual explanation. Specifically, the point in this line of work is that there are cases in which understanding does not imply truth and fact. Therefore, Kvanvig will focus on the cases of the factuality of understanding to the extent that they are most useful for the theoretical project of epistemology, and of course in education.

There would be two fundamental uses in this perspective, the objectual use and the propositional use. When understanding is attributed to a propositional operator by understanding that something is the case; and when the understanding is grammatically followed by an object such as politics or gravity. The factuality in the latter case is not direct insofar as the truth and falsehood occur to the positions. Beliefs about an object must be true. How to distinguish between knowing and understanding Colombian politics?

But if when propositional understanding and understanding of objects are attributed it implies having

knowledge of both propositions, what is it that the understanding adds that knowledge lacks?

Kvanvig's proposal is that "the central feature of understanding lies in the neighborhood of what coherent theories say about justification" [6]. Let us remember that coherentism is a theory of justification epistemic of knowledge, which faces the lineal and hierarchical nature of beliefs as proposed by foundationism, where there are basic or fundamental beliefs that when in contact with the experience and the data of the sensory experience of which they derive the rest of the beliefs by inference. The image of a pyramid or a building illustrates this position if we accept that the fundamental beliefs are at the base, and it is from that structure that the rest of the theoretical system is assembled, thus solving the justification.

For coherentism the justification is not linear but multidirectional. That is, the justification is not established towards individual propositions but towards the set of beliefs that make up the system. For coherentism beliefs support each other in a variety of ways in ways that are mutually supportive. According to Rana Rosaleny "the supporting structure of the entire belief system is thus not a chain but - hence the importance of images or metaphors in coherentism - a network, an arch of bricks or, for example, a raft (To quote the famous image of Neurath" (Rana Rosaleny, [11]) In this way, beliefs are justified in their "being coherent with a comprehensive belief system", as Rana Rosaleny puts it.

However, if the justification proposal for coherentism is based not on basic beliefs but on the relationships of mutual support between specific beliefs and the rest of the cognitive system, then this proposal can be questioned of circularity to the extent that it is self-justifying. There have been many exits to this accusation in which I will not go into detail, however, it must be mentioned in passing that one of its most outstanding features is consistency, that is, the non-contradiction between the elements that constitute the system, and the Requirement of connections inferences that has to do with the idea that there are inferences that fulfill explanatory functions with respect to other beliefs, in this way we have then that the coherence of the system will be of a greater degree as there are more mutual explanatory relationships established. "A belief system is better justified than another, if it has a greater degree of coherence" [11].

Finally, some authors such as Susan Haack [5] accept the existence of observational beliefs that are somehow independent of other beliefs, since since Kant (Ferraris, [3]) experience has played an important role in the formation of beliefs; some perceptual beliefs are those that connect with the truth. To admit that there are data that come from perception and others that do not is not to postulate an asymmetry since the permanence of beliefs in the system is not determined by the origin (whether it is perceptual or not) but rather the increase or decrease of coherence. Indeed, it should not be forgotten that the strength and durability of beliefs will be greater if there is an empirical basis to support them. Experience intervenes in the justification only as an additional requirement, but it is not preponderant.

However, as Rana Rosaleny says, this position is still problematic since, "(...) it is possible to conceive a coherent set of propositions around a fantasy world that exhibits great continuity and systematic stability, without the derived beliefs of that fictionalized world have any correspondence with reality" [11]. Which is dissolved if we consider that perceptual beliefs do not lack theory, they are not isolated but are traversed by our conceptual schemes. Let us leave this brief digression about coherentism here, as it takes us far from our subject if we continue to consider in detail its strengths, criticisms, and disadvantages. Let's move on properly to the notion of understanding.

If knowledge consists of loose fragments of information, such as propositions, once these fragments are put together we are in the realms of understanding. Understanding requires the capture of logical, explanatory, probabilistic relationships, and these exercise the role of justification. The use of coherentism in the theory of understanding has, in Kvanvig's opinion, the advantage that by not being based on a singular proposition, but rather on a larger body of broad information, understanding attributions are given groups and subgroups of beliefs, which allows assessing inadequate theories in relation to traditional concepts of knowledge and justification. Hence, he maintains that the justification comes in degrees since two bodies of information differ from each other by their degrees of coherence.

On the question of whether understanding is a form of knowledge, the answer is no; and the reason for this is that what is central for knowledge are the connections between mind and world, while for understanding what is central are the captures of relationships and combinations of information fragments, therefore there is no logical connection between the two. He will say that, "what is crucial for understanding is to see or appreciate internally the explanatory and coherence-inducing relationships that exist in a body of information" (Kvanvig, [6])

Thus, for example, there are inconsistent theories but that we can understand as the case of set theory, in which its axioms and theorems that follow from these can be grasped, without the commitment of those statements to the truth. A person with false beliefs about a body of knowledge may have an understanding of it, since the false beliefs would only be peripheral, so there would be a percentage of true beliefs. The value of understanding for Kvanvig is focused on "capturing relationships of coherence. Such coherence relationships in this context contribute to justification. This justification is subjective" [6]. The person is the one who captures and intuitively the pieces of information to put them together, however, as the understanding has factual notes that captured by the person must be correct.

Finally, the answer that Kvanvig to the question about the value of understanding is that what is valuable in itself are explanatory relationships since it leads to discovering new truths, but mainly because it gives order and systematicity to thought when it is directed to an object of study.; Something that simply by adding true, even justified beliefs, would be far from attain. The most interesting thing about the

epistemological approach based on understanding is that the organization of thought allows reasoning through the different fields of information, which puts it at the forefront of transdisciplinarity, which has a utility that goes beyond the theoretical, as it becomes a platform from which the action unfolds. No less important is that we can attribute a feeling of satisfaction and fulfillment to the person when at the end or during the research process an understanding of an object of study is reached. This perspective is shared by Strevens [12].

#### 4. Understanding, Fiction, and Thought Experiments in Science

In 2006 Catherine Z. Elgin published "From Knowledge to Understanding" in a compilation entitled "Epistemology Futures", in Oxford. In this article he draws as an outstanding feature of scientific activity the act of representing and thinking about an area, if what we want is to understand, but understanding implies selecting, manipulating and devising both the data and the representations that one has of them. The representation depends on the classification or categorization that we make of the individuals or classes that constitute a field of knowledge. Hence, the cognitive interests of a given science depend on the taxonomy or the category scheme that it deploys.

What allows us to affirm that some things are superficially similar or different is the ability of categorization. Categorization allows selection through criteria or filters, and this is provided by understanding. Similarities and differences could be overlooked if not because science postulates what to focus on. In Elgin's words, "We conclude that rabbits or hares are or are not the same type of animal. In such cases we fail to understand the phenomenon, although our explanation is made up of justified true beliefs" [2].

Science has specific objectives and representation is what gives it the ability to focus on what is relevant, so in the particular cases that are the object of study, vagueness cannot be eliminated for the sake of the defined and sharp lines of knowledge. There are issues that are considered central, but later elaborations are introduced with the sole purpose of defining the phenomenon in question, delimiting it according to the problem to be solved in mind. That is why Elgin proposes to speak of Focus of representation to highlight explaining how we highlight what is really important about the peripheral. The scale, scope and content of the representations are always susceptible to change as new purposes emerge to the detriment of the original purposes of the representation. The classes, their degree of generality and how they should be represented is conferred by the understanding of a specific domain scope.

As thought experiments are not replicas of phenomena, their cognitive contribution to science depends on the interpretation we make of them, and that interpretation depends on some underlying assumptions, a thesis shared by Brendel, as we will see later. Thus, for example, for Brendel [1], if the thought experiment is reduced or translated to the

argumentative or narrative form (a thesis taken from Norton), the presence of hidden premises that function as background knowledge of which no we need to be aware. Mental experiments for Norton [8] are ordinary arguments disguised as pictorial or narrative forms. As a self-avowed empiricist he holds that the raw material of mental experiments is the knowledge we already have of the world, whether explicitly or tacitly. In that sense, the limit of mental experiments seems to coincide with that of argumentation. Consequently, a key point in Norton's position is that mental experiments are determined by logical rules, whether inductive or deductive.

Back to Brendel the characteristic of this background knowledge is that they are understood as intuitions expressed in propositional attitudes accompanied by a feeling of certainty. The most important thing for this author is that part of this intuition comes from common sense, and from background scientific knowledge. The evaluation of knowledge in thought experiments then depends on intuitive knowledge. This trait is what makes the experiment representative of nature. Hence the experiments are quite artificial. This is why it is currently unanimously accepted that science distances itself from phenomena by creating models, idealizations and thought experiments.

Following Elgin then we accept that scientific models are schematic representations of features considered relevant and that they obviously leave out those that are not. In this way the relevant features are amplified with the important purpose of highlighting the specific consequences. Elgin even goes so far as to say that the models do not describe anything about the world since they operate for certain purposes. Models could be said to describe things that don't happen in the world. The recurring example is the gas model which represents the molecules as if they were perfectly elastic, dimensionless and that they do not show mutual attraction. The model is idealized because it focuses on features such as temperature, pressure, and volume that are critical to understanding actual features.

Mental experiments in science are intentional representations of circumstances or events whose purpose is to validate, deepen and contrast theoretical explanations through the construction of hypothetical scenarios; However, that would not be their only purpose as they are also used in science teaching in terms of modeling in science learning; in both cases, conjectures and possible consequences are presented or artificially created. These guesses have a theoretical platform in which variables are assigned in order to manipulate and control data changes. Mental experiments usually allow us to go beyond experimentation in the real world, and in terms of didactics they facilitate understanding since they explain and clarify the abstract states of things. The ability to manipulate and control changes in mental experiments give the trait of flexibility, since these can be rethought and thus carry out different versions of the same scenario and generate possible postulates (Yirsén Aguilar, Ángel E. Romero, [10]). For example, for Lukás Bielik, a professor at the University of Bratislava, thought experiments

appear as evidence for or against some philosophical and scientific theses or theories.

The mental experiment must be considered as an action of thought in which remembered experiences are valued and language as long as it fulfills the narrative function of public presentation of the same; thus, from the imagination you can create or visualize possible worlds. In the case of the explanation of a theory or concept through a mental experiment, it would be necessary to ask what kind of characteristic and function would it have? That is, to delve into the theory, substitute the possibility of factually experimenting, explaining and clarifying the activity of consciousness? What kinds of theories and hypotheses can be evaluated, how much intuitive knowledge is used, how many contradictions between theories is put.

The most important thing is that thought experiments have little to do with knowledge and more to do with understanding phenomena. Even though the imagined conditions will never happen, the experiments have high degrees of effectiveness. Thus, for example (Elgin, [2] considering a person going up in an elevator with and without the presence of a gravitational field, Einstein showed the equivalence between gravitational and inertial mass, just as Galileo discredited the Aristotelian thesis of motion, while discovering that the acceleration with which bodies fall is independent of their weight.

Does the fact that both experiments are not faithful to the exact considerations of nature discredit them? These so-called realistic anomalies in theories neither do nor make them faulty. Models, idealizations, and thought experiments are theoretical devices that contribute to science effectively. The inevitable question then is how do thought experiments contribute to understanding?

The idea that thought experiments have to do with unreal situations, counterfactual assumptions, fictitious and idealized is almost an agreement. Among the functions that experiments have is to identify theoretical contradictions in a domain of knowledge, support a concept or a theory, illustrate an abstract theory and detect conceptual vagueness. Do these features lead to understanding in its most general sense?

For Elgin as well as for Ornelas in thought experiments imagination or fiction play an important role. Fictions have the advantage of being exempt of the truth conditions; However, contrary to general common sense, they are not merely speculations with high degrees of freedom, as they have constrictions; Thus, for example, while for Ornelas [9] the imagination has empirical notes that relate it to the truth, given that "imagining something is presenting us with a candidate for knowledge, a prediction with a high probability of being true" Ornelas, [9]. Imagination is defined in this way as a cognitive capacity that allows to create in a counterfactual way, new mental representations, from previous stored experience; which ranks it as a by-product of perception and whose purpose is to anticipate dangers and learn about the characteristics of the environment. The imagination then produces reliable beliefs.

To answer the same question, that is, how does a fictitious

representation, not a true one, understand something of the world? He will say that it does so "by exemplifying features that diverge (at most) in a negligible way from the phenomena with which it deals," [9]. Thus, for example, following Elgin, Einstein dispenses with many drawbacks when he imagines a person traveling at the speed of light: that whoever travels at that speed will acquire an infinite mass, which he could not see because his retina would be reduced to the size of a light photon, and many more inconveniences that are considered insignificant in the context of the thought experiment. There are some background assumptions that will exert constraints on the design and interpretation of the experiment. This is what allows some traits to be maintained while others are disregarded.

Elgin suggests that the same applies to scientific theories understood as models. Both the models and the thought experiments have differences with respect to the phenomena that are insignificant; if the differences are not negligible, you run the risk of doing bad science. Determining what is insignificant or negligible of what is not requires the focus of attention. The models can be refined to achieve a better correspondence with the facts. Sometimes the simpler models are more revealing.

Fictions in science are very important from the cognitive point of view, and a theory that it holds should not be considered pejorative. He states that points of mass in fields of gravity are easier to conceptualize than considering planets with their actual dimensions. These fictitious approaches then take the limelight away from the traditional criterion of truth. Truth was understood as the relationship between the content of a proposition and the world, but if this conception is too short to think about scientific theories, are we opening the door to relativism and pseudoscience? Elgin will say that it is not necessary since holism in the style of Quine does not neglect the role played by the court of experience; empirical observations discredit or confirm scientific theories.

Another very important point of view of Elgin is that these cognitive devices, namely models and thought experiments, play a causal role insofar as they allow scientists to understand how things are, then veritism is not absolutely discredited, since that the role of these devices is not constitutive of knowledge but rather embodies understanding.

The understanding that mental experiments provide from the perspective of the imagination makes it possible to generate conclusions in any field of research; these conclusions can be expressed with what Ornelas, following Williamson, calls counterfactual reasoning: what they do is subsume observed cases and cases to come. This reasoning requires a conditional structure of the form (If X then Y), to then establish the form (If X were such a thing) and the consequent was logical (then Y would be such a thing): this would be the procedural character of the narrative of thought experiments. And of course, the cognitive foundation of counterfactual reasoning is imagination. But it is not a freedom in its products, it is not unrestricted, as there are factual restrictions to prevent anything from being covered.

In any case, for Elgin, understanding is what allows establishing within a scientific theory the reflective balance between fictions, methods, empirical statements, categories in terms of internal coherence, as we saw with Kvanvig. Finally, Elgin affirms that literal and factual truths lose relevance in favor of the complex symbolization that only good science exhibits. And he closes his article by saying that "Understanding a domain in terms of a theory is being in a position to recognize, reason, anticipate, explain and act on what happens in the domain based on the resources that the theory provides" [2].

## 5. Final Considerations on Thought Experiments

Thought experiments that play the above roles can improve the quality of our knowledge of epistemological relations to the world without necessarily (or simply) increasing our stock of justified true beliefs. Stuart [13]. Some of the basic characteristics of thought experiments is that the purposes of the mental experimenter can be achieved without the need to perform the actual experiment. That is, their results, even if they are not perceptible, can be "seen" by means of a certain intuitive capture.

The counterfactual and idealized feature of some of the thought experiments such as Maxwell's demon, Einstein's observer who travels mounted on a ray of light, or Galileo's bodies (this experiment has been carried out in vacuum chambers on Earth and in the Moon, Apollo 15 mission, in 1971) that fall without air resistance, are examples of thought experiments that in principle could not be carried out, in the case of Galileo, later they were. However, there would be other thought experiments that definitely cannot be carried out like Newton's cube since this experiment requires as one of its assumptions to be carried out in a completely empty universe, therefore it is not possible to carry it out, and the train of Einstein in which a possible or real situation is imagined. Brendel points out some of the most important characteristic functions assigned to them, namely:

That mental experiments have the function of refuting basic statements that constitute narratives about possible worlds. In this sense, the mental experiments would aim to locate hidden contradictions when constructing hypothetical scenarios that problematize traditional conceptual structures. Scientific progress depends largely on the detection of inconsistencies in scientific theories.

However, there is another more constructive purpose such as providing evidence in favor of a theory that has been questioned, for example, Newton's bucket experiment was built to demonstrate the absolute character of space. In this positive sense, thought experiments can also assume a pedagogical function, which in my case is what interests me. This implies the understanding of very technical theories due to its high degree of abstraction. The example presented by Brendel is that of the prince and the shoemaker by John Locke to explain psychological continuity as necessary to

think of personal identity as rational selfhood.

Provisionally, thought experiments achieve their purpose without the need for their actual execution. They share with real experiments the possibility of modifying data, by showing the functional dependence between variables in a constructed scenario. Sometimes they prove that certain concepts are contradictory. In other cases what they do is provide evidence or support for a theory. Pedagogically, they illustrate an abstract or complex position, and finally, they detect conceptual vagueness. Before moving on to the Nersessian thesis, we will refer to some classic positions of thought experiments.

Mental experiments are exemplifications of situations that depend on the range of phenomena to be highlighted and therefore on the purpose of the researcher or the teacher, and their purpose is to capture ideas and gain understanding. In other words, for Nersessian [7] mental experiments allow simulating an exemplary or representative situation in order to detect the implications derived from the representation of a specific phenomenon or event. Building models to make inferences using cognitive mechanisms is a great way to make predictions, but also to teach science, as we'll see later. Can science be taught using thought experiments? Can modeling be taught, that is, to build thought experiments? The answer to both questions is yes.

---

## References

- [1] BRENDEL, E. (2003): "Pompas de intuición y el uso adecuado de los experimentos mentales", en *Ideas y Valores*, N. 123, Diciembre de 2003, Bogotá Colombia.
- [2] ELGIN, K. (2011): "Del conocimiento al entendimiento", en *Normas, valores y virtudes epistémicos. Ensayos de epistemología contemporánea*, Margarita Valdés y M. A. Fernández (compiladores), Universidad Nacional Autónoma de México.
- [3] FERRARIS, M. (2007): *Goodbye, Kant! Que queda hoy de la Crítica de la razón pura*, Editorial Losada.
- [4] FRICKER, M. (2011): "El valor del conocimiento y la prueba del tiempo", en *Normas, valores y virtudes epistémicos. Ensayos de epistemología contemporánea*, Margarita Valdés y M. A. Fernández (compiladores), Universidad Nacional Autónoma de México.
- [5] HAACK, S. (2012): "La unidad de la verdad y la pluralidad de las verdades", en *Teorías contemporáneas de la verdad*, Tecnos.
- [6] KVANVIG, J. (2011): "Conocimiento y entendimiento", en *Normas, valores y virtudes epistémicos. Ensayos de epistemología contemporánea*, Margarita Valdés y M. A. Fernández (compiladores), Universidad Nacional Autónoma de México.
- [7] NERCESSIAN, N. (2018): "En el laboratorio del teórico: la experimentación mental como construcción de modelos mentales", en *Trabajando en el laboratorio de la mente: naturaleza y alcance de los experimentos mentales*, UASLP Universidad Autónoma de San Luis Potosí, Facultad de ciencias sociales y Humanidades, México.

- [8] NORTON, J. (2018): "Por qué los experimentos mentales no trascienden el empirismo", en *Trabajando en el laboratorio de la mente: naturaleza y alcance de los experimentos mentales*, UASLP Universidad Autónoma de San Luis Potosí, Facultad de ciencias sociales y Humanidades, México.
- [9] ORNELAS, J. (2016): "Un estudio metafilosófico de la metodología científica", en *Praxis filosófica. Nueva serie*, N. 43, Universidad del Valle.
- [10] ROMERO-CHACÓN, A. E. (2017): *La experimentación en la clase de ciencias. Aportes a una enseñanza de las ciencias contextualizada con reflexiones metacientíficas*, editorial Universidad de Antioquia.
- [11] ROSALENY, R. (2017): *Problemas de la teoría del conocimiento. Una introducción a la epistemología contemporánea*, Universidad de Antioquia.
- [12] STREVEN, M. (2017): "How Idealizations Provide Understanding", in *Explaining Understanding. New perspectives from Epistemology and Philosophy of Science*, Edited by Stephen R. Grimm, C. Baumberger and S. Ammon, Routledge, New York.
- [13] STUART, M. (2018): "How thought experiments increase understanding", in *The Routledge Companion to Thought Experiments*, Edited by T. Stuart, Y. Fehige and J. Brown, Routledge Philosophy companions, 2018.
- [14] TUFFANELLI, L. (2010): "La comprensión como problema", en *Comprender. Qué es? cómo funciona?* ediciones Narcea.
- [15] VILLORO, L. (2006): *Crear, conocer, saber*, Siglo XXI editores.
- [16] ZAGZEBSKI, L., (2011): "La búsqueda de la fuente del valor epistémico", en *Normas, valores y virtudes epistémicos. Ensayos de epistemología contemporánea*, Margarita Valdés y M. A. Fernández (compiladores), Universidad Nacional Autónoma de México.