

THOUGHT EXPERIMENTS AND DISGUISED ARGUMENTS

Juliano Santos do Carmo
Universidade Federal de Pelotas

“Imagination and its shared products play a crucial role in human social life. Almost all important works of fiction transmit a “message” or “messages” that are communicated by the text but are not in the text”¹.

Abstract: One of the most controversial methods of philosophy is the use of thought experiments. If, on the one hand, philosophers usually agree that a hypothetical reasoning or an imaginative example can refute a previously well-supported theory or hypothesis, on the other there is a strong skepticism about the use of thought experiments in Philosophy. The skepticism is reinforced by the fact that scientific thought experiments are profoundly influenced by the expectation that the outcome of a thought experiment may reflect the outcome of an actual experiment, whereas in Philosophy the use of this tool does not necessarily reflect what happens in the world. In other words, scientific thought experiments seem to have well delimited criteria, while philosophical thought experiments do not. In this work I will argue in favor of the thesis that not all thought experiments are disguised arguments, but that nevertheless they should not be considered as lacking in criteria of correction.

Keywords: Thought Experiments; Intuitions; Philosophy of Philosophy.

Introduction

One of the most controversial methods of philosophy is the use of thought experiments. If, on the one hand, philosophers usually agree that a hypothetical reasoning or an imaginative example can refute a previously well-

¹ SEARLE 1975, p.319-332.

supported theory or hypothesis², on the other there is a strong skepticism about the use of thought experiments in Philosophy. The skepticism is reinforced by the fact that scientific thought experiments are profoundly influenced by the expectation that the outcome of a thought experiment may reflect the outcome of an actual experiment, whereas in Philosophy the use of this tool does not necessarily reflect what happens in the world. In other words, scientific thought experiments seem to have well delimited criteria³, while philosophical thought experiments do not⁴.

In face of this skepticism, some philosophers have sought to establish certain criteria of validity for philosophical thought experiments. One of the most influential strategies was presented by John Norton⁵, who's the central idea was that thought experiments are "disguised arguments". In this sense, all thought experiments should be converted into arguments so that, with the help of our usual rules of inference, to ascertain whether thought experiments are useful to lead us to true conclusions from some imagined scenario. Although this perspective has gained a great deal of attention from the philosophical community, I think that the conversion or translation suggested by Norton would bring about a great loss in terms of the persuasive force that many of philosophical thought experiments possess. So, I will argue in this paper that not all thought experiments can be converted into arguments and that this fact does not imply that there are no criteria to evaluate them.

1. What are thought experiments?

Thought experiments perhaps belong to what we might call of "hypothetical reasoning" (alongside hypothetical-descriptive examples, counterfactuals, etc.). Hypothetical reasoning are ways of reasoning or cognitive activities in which we held assumptions and we try to draw its consequences. In other words, they are ways of reasoning about alternative worlds regardless of their similarities to the real world. It is supposed that hypothetical reasoning can provide evidence for or against a given theory or hypothesis, through the test of their logical consequences.

² One of the most important example is the Edmund Gettier's use of thought experiments to refute the traditional analysis of knowledge as justified true belief (1963).

³ Most scientific thought experiments reflect real possibilities, and in this sense all variables need to be controlled as in a real experiment. See: BRENDEL 2004.

⁴ In fact, some philosophers thought that philosophical thought experiments are profoundly unlike those in natural science. I don't believe the differences are so drastic, since the imagination plays a vital role in the evaluation of counterfactual conditionals.

⁵ NORTON 1994, pp. 333-366.

Ernst Mach (1838-1916) was one of the first authors to use the term “thought experiment”⁶. In *Zeitschrift für den physikalischen und chemischen Unterricht* (1897)⁷, Mach defends that there are many kinds of thought experiments, such as “dreaming”, “writing novels”, “describing imagined situations”, etc. Mach claims that this tool had also a fundamental role in natural sciences, once each empirical investigation should start necessarily with one thought experiment. Scientists should begin their studies with one hypothesis in which all variables are controlled before engaging into a real experiment. This idea is very common in theories of contemporary philosophers of science. However, this characterization does not seem to provide us with criteria for evaluating philosophical thought experiments.

A central difference between thought experiments and real experiments (in philosophy and natural sciences) is that in the latter the scope of possibilities is limited to the actual world, whereas philosophical thought experiments have a much wider scope of possibilities: thought experiments in this field can appeal to any possible or even impossible scenarios. Thus, some thought experiments are free from the demand and limits of practical realizability and empirical testability⁸. Of course, some thought experiments cannot be realized once they make use of unreal or counterfactual situations, while others can easily be converted into real experiments. There are at least two kinds of thought experiments: those that involve analytical reasoning and those that involving synthetic reasoning. In most cases, we can assume that the second kind can be converted into real experiments while the first not. But the differences are not limited to these.

Some authors, like John Norton (2004)⁹, believe that all thought experiments are analytical tools and, in these sense, they can all be converted into arguments. Otherwise, according these authors, there would be no “a criterion of rightness” for the use of thought experiments. Once converted into arguments, we could easily evaluate them in terms of their logical validity, consistency, and so on. As with the arguments, some thought experiments would be considered as solid arguments (with true premises and conclusion, and in the form in which the conclusion is logically bound or supported by the premises), while others would be considered as bad arguments insofar as they

⁶ GENDLER 1998, 2004.

⁷ MACH 1897, pp. 1-5.

⁸ LEHTONEN 2012, p. 5.

⁹ John Norton has been primarily concerned with scientific thought experiments, but his account can be easily extended to philosophical thought experiments.

fail in satisfying our usual inferential rules¹⁰. About the argumentative function of thought experiments, their job is to support or refute claims. It has even been claimed that all thought experiments are arguments, that is, “wholes” consisting of premises, background assumptions and conclusions¹¹. Accordingly, the value of thought experiments depends on their logical validity, credibility and persuasiveness.

Although it seems to be a useful criterion for evaluating thought experiments, it excludes many other kinds of thought experiments that are not merely testing the scope of concepts, testing theories or hypotheses for contradictions, illustrating or illuminating one complex or abstract position, detecting the vagueness or border cases of a concept (and all that supposedly an analytical thought experiment presupposes). In other words, this criterion excludes all thought experiments involving synthetical reasoning.

One of the most important characteristics of second kind of thought experiments is the fact that they can provide new information about the natural world. Of course, this is a very controversial claim. In fact, someone could ask: How hypothetical reasoning could provide us with new information about the natural world? Where do they come? This problem has been called, by Norton and others, the “Informativeness Problem of Thought Experiments”. Well, in my opinion, the answer of this problem depends on how we understand of role of hypothetical reasoning in our philosophical investigations. One hypothesis is that the new information is found by rearranging the data that were empirically obtained and by the underlying logical reasoning. This idea is defended by Michael Bishop when he defines thought experiments as mental representations of experiments in which the results are derived by a process of reasoning.

The reasoning used in thought experiments is not merely analytic, but also synthetic since it employs substantive assumptions about how the world works. The assumptions can include, for example, laws of nature, causality, and psychological laws¹².

Elke Brendel (2004) also defends that thought experiments can offer us new information about the world, even though they are based on no new

¹⁰ It may be important to note that our best-established rules of inference can also be considered as human institutions. Nelson Goodman makes this very clear in *Fact, Fiction and Forecast* (1965, p. 66): “a rule is amended if it yields an inference we unwilling to accept, an inference is rejected if it violates a rule we are unwilling to amend”.

¹¹ NORTON 1996, p. 354.

¹² BISHOP 1998, p. 20.

empirical data. Brendel claims that the thought experiments, like real experiments, also study “the dependency of variables on other variables and the dependency of actions on other actions. In this sense, thought experiments also depend on some background assumptions or background theory”¹³. In addition, Brendel also seeks to show that certain thought experiments aim to investigate our “intuitions” about the world and our social relations. Unlike James Robert Brown¹⁴ who defends a Platonic view of intuitions (as special windows for the world of the laws of nature), Brendel offers a more naturalized perspective of intuitions. According to Brendel, intuitions are mediated by logical reasoning about knowledge that we already possess. We gain new information of intuitions that emerge from thought experiments “by rearranging or reorganizing already known empirical data in a new way and drawing new inferences from them or by looking at these data from a different and unusual perspective”¹⁵.

Tamar Szabó Gendler (1998) argued that thought experiments can teach us something new about the reality by helping us to reconceptualize the reality in a better way. Thus, Gendler gives thought experiments an important role, specially in the development of new conceptual tools. Similarly, Thomas Kuhn (1977) defends that thought experiments should be recognized as a powerful tool to increase the understanding of our conceptual system. He believes that each successful thought experiment personifies in its structure some prior information about the world. This information is not discussed in the experiment; rather, it rests on empirical data, which must be well known and accepted, usually before the experiment is conceived. In other words, nothing about the hypothetical situation may be entirely unfamiliar or strange.

Recently, Timothy Williamson (2007) has advocated a perspective in which philosophical thought experiments can be converted into arguments, but according to Williamson, the imagination plays an irreducible role in warranting the premises. It seems Williamson is interested in the arguments that “may arise” from or which “underlie” philosophical thought experiments, and this strategy is quite different from that was advocated by Norton (his interest is focused on the literal conversion of thought experiments into arguments). It should be noted that Williamson also defends an important role for counterfactual conditionals in our usual modes of reasoning. In his non-epistemic approach about the Gettier’s cases¹⁶, paradigmatic philosophical

¹³ Cf. BRENDEL 2004, p. 91.

¹⁴ BROWN 1991a.

¹⁵ BRENDEL 2004, p. 95.

¹⁶ GETTIER 1963, pp. 121-123.

thought experiments are simply valid arguments about counterfactual possibilities¹⁷.

Maybe a first step in characterizing the different kinds of thought experiments could be to differentiate them from scientific thought experiments. Which could, in principle, easily be converted into arguments. A second step could be to classify them into analytical and synthetical, since the former does not expand our knowledge while the latter supposedly do. On the other hand, we can to classify thought experiments that cannot be converted into arguments or that are not formally arguments like “intuition pumps”. This term has been used by Daniel Dennett to classify the different types of intuitions that emerge from thought experiments. Some are legitimate intuition pumps while others are not. I will say something about this strategy at the end of this paper.

However, if all of thought experiments were converted into arguments, as Norton and his colleagues defend, then many forms of thought experiments in philosophy should be recognized as bad arguments. Possibly, all forms of synthetical thought experiments cannot be translated into arguments since most of them do not have conclusive premises or ending with interrogative sentences. Moreover, in many cases the translation of the thought experiment into the argumentative form fails to capture the target for which the experiment was designed. I will discuss this problem in more detail in the next section.

2. Thought Experiments and Disguised Arguments

According to Norton, thought experiments are fully eliminable and their persuasive force is the result of their being sound disguised arguments. In this sense, the justificatory force of thought experiments comes from the fact that it can be reconstructed as an argument with explicit premises that make no reference to imaginary particularities that could be affect the generality of the conclusion of this argument. In other words, thought experiments lead us to justified conclusions because they are disguised arguments. His “elimination thesis” presupposes that any conclusion reached by a successful thought experiment will also be demonstrate by an argument that is not a thought experiment¹⁸.

It’s important to note that there are deeply differences between “thought experiments” and “arguments that are not thought experiments”.

¹⁷ WILLIAMSON 2007, p. 207.

¹⁸ According to Gendler, this thesis is ambiguous: it is compatible with both a weaker reading, which he called the “Dispensability Thesis”, and a stronger reading, which he called the “Derivativity Thesis”.

Firstly, to draw a conclusion of a thought experiment can be thought like make a judgment about what would happen if the particular state of affairs described in some imaginary scenario could be obtained in a real experiment. Secondly, to draw a conclusion of an argument that is not a thought experiment is to realize a process of inductive or deductive reasoning from a set of explicit premises that no refer to hypothetical or counterfactual states of affairs that could be affect the generality of the conclusion. This characterization suggested by Gendler implies that thought experiments differ from arguments that are not thought experiments in two crucial aspects:

(i) Thought experiments are not presented as arguments, but rather *as invitations* to contemplate possible state of affairs.

(ii) Thought experiments make essential reference to hypothetical or counterfactual states of affairs¹⁹.

Norton agree with the idea that an essential characteristic of thought experiments is precisely the appeal to hypothetical or counterfactual situations (otherwise they would be real experiments), but he also defends the idea that any good thought experiment can be translated into an argument without loss the persuasive force. That is, the elimination thesis involves a process of argumentative reconstruction in which the fictional discourse is replaced by explicit premises that can be evaluated in terms of our usual inferential rules. At the end of this process, explicit premises do not refer to hypothetical or counterfactual situations presented in the imagined scenario. Norton claims that this process does not affect the persuasive force of thought experiments.

We can assume that one thought experiment is a process of reasoning carried out within the context of a well-articulated imaginary scenario to answer a specific question about a non-imaginary situation. Gendler defends the idea that both the powers and limits of this methodology can be traced to the fact that when the contemplation of an imaginary scenario brings us to new knowledge, it does so by forcing us to make exceptional cases meaningful²⁰. In Gendler's opinion, some thought experiments can provide us with new scientific knowledge through a careful examination of the guided contemplation of an imaginary scenario. And this is something that an argument alone cannot to do.

In this sense, whether an imaginary case cannot be differentiated from a real case, our responses to imaginary case are as relevant as our responses to a real case would be. So, for example, there seems to be no difference between deciding whether an actual agent, like Phineas Gage

¹⁹ GENDLER 2000, p. 35.

²⁰ GENDLER 2000, pp. 15-22.

(described by Antônio Damásio in *The Descartes's Error*), cannot be blamed for his supposed immoral deeds, and deciding whether, if there were an agent with damages in the frontal lobule, that agent cannot be blamed for his supposed immoral deeds. But to the extent that hypothetical cases allow idealizations and real cases do not, then it is possible that both cases are completely different. Just note that we don't need looking at the world to search actual states of affairs that corresponds to the case described by the hypothetical case.

There are some problems with the idea that any thought experiment can be reconstructed or translated as an argument that is not a thought experiment. We can summarize some of these problems as follows:

(1) Thought Experiments invite us for a constructive participation, while the arguments do not play a constructive role (they are deductive or inductive).

(2) Thought Experiments make essential reference to hypothetical or counterfactual situations and it is not always possible to eliminate the particularities described without causing the loss of its persuasive force.

(3) Thought Experiments describe imaginary scenarios to answer questions that are not merely hypothetical or imaginary. Arguments are usually used to validate theses or hypotheses.

(4) Thought Experiments describe particularities in ways that can make manifest our background knowledge, because sometimes they track intuitions of our conceptual system. Arguments do not track intuitions, but only demonstrate valid forms of reasoning.

(5) Thought Experiments sometimes can be converted into arguments that are not thought experiments. There seems to be no explanatory gain in translating arguments into thought experiments.

(6) Thought Experiments can be both analytical and synthetical tools. Arguments do not provide new information about the world, but they test the validity of our assumptions about it.

Of course, this summarized list is far from exhaustive and some of these problems are more serious than others, so I will focus only on cases (2) and (4).

2.1 The Persuasive Force of Thought Experiments

Thought experiments can change our views about certain concepts and help us decide about what we should say or do in certain circumstances. This happens not so much by the logical form in which the hypothetical reasoning is presented, but mainly by the constructive movement that we must engage to make significant the exceptional cases present in the experiment. In

this sense, if we eliminate the particularities (hypothetical or counterfactual situations) of the imaginary situation, we also isolate it from a network of events of which it is a part. In doing so, we eliminate the persuasive force of thought experiments. Thought experiments often present non-empirical assumptions and have a discursive flavor, which the Norton's view fails to capture. In the extent that the Norton's view to exclude the particularities presents in hypothetical scenarios, thought experiments cannot have persuasive force. I think that the Norton's view is wrong, because the persuasive force is just in the particularities involved in thought experiments.

According to Gendler (2004), the function of thought experiments is also to show how our conceptual systems can be modified or maintained and therein lies their persuasiveness. According to Norton (1994), the persuasiveness of thought experiments lies in their logical force when they are reconstructed into arguments. In the constructivist view of Gendler²¹, the persuasiveness of thought experiments also lies in their discursive force. In the extent that the Norton's view eliminates the particularities that are present in the imaginary scenarios, she also eliminates their discursive force and, for consequence, their persuasive force.

Evidently, thought experiments often cannot be real experiments because they rely on hypothetical or counterfactual reasoning and employ abstraction and idealization for reach their results. They cannot be Platonic devices, because they don't need presupposes a special kind of perception to reach their outcomes and they don't cannot capture relations between universals (like Brown claims). Not all thought experiments can be translated into an argumentative form (deductive or inductive), because intuitions and imagination can play an important role in these forms of reasoning. The most important role of thought experiments is providing *understanding*.

2.2 The role of thought experiments is providing understanding

Thought experiments can be used to providing understanding about certain philosophical issues. This means that the evaluation of the outcome of a hypothetical case can go beyond what is described in the imagined scenario. It is easy to observe this in most philosophical thought experiments (in all areas of philosophical interest). Hubert Schwyzer (1969), for example, elaborated an influential thought experiment to show that the rules of chess are not sufficient conditions to playing chess. The experiment was as follows:

²¹ GENDLER 1998, p. 420.

Imagine the following: I have now been in Ruritania some time, and can speak the language tolerably well. One day I accompany my host with his family to a certain building, where many people are sitting in a circle on the floor, murmuring among themselves agitatedly. In the center of the circle is a small table with two chairs, and on the table, is a chessboard with chessmen arranged as for the beginning of a game. After a while, two men in elaborate clothes enter the room and seat themselves at opposite sides of the table; whereupon those sitting on the floor fall silent and watch intently. The men at the table then proceed, with what appears to be an air of great concentration, to move the chess pieces around on the board according to the rules of chess. It strikes me, however, that they play a rather wild game, and I can see no consistent strategy in the moves of either player. The excitement mounts until, after an hour or so, white mates black. Then everyone present, including both men at the table, shows signs of extreme relief; they mop their brows, smile and congratulate one another²².

The outcome of this thought experiment is precisely the idea that in “Ruritania’s chess game” there is no winning or losing, because chess is a sacred rite (chess is played once every year by the priest of community, for purposes of determining the will of the gods. If white mates black, the community and the crops will flourish; if black mates white, there will be trouble). Note that the Schwyzer’s example also could be perfectly observable in the real world (there is at least a logical possibility). But if we want to *understand* how this case could teach us something about chess (and about any other competitive game), we need to think about the nature of the constitutive rules of chess and their dependency of regulative rules (or meta-institutional concepts). Seems clear that a competitive game like chess is dependent of our meta-institutional concepts like “victory” or “defeat”, for example, but this does not seem to follow immediately from the argumentative form of this imaginary case.

When we translate the Schwyzer’s example into an argument, however, we observe that (evaluating the hypothetical situation and eliminating the particularities that could affect the generality of the conclusion) still does not seem easy to see how the conclusion in favor of meta-institutional concepts follows exclusively from the explicit premises. In the scope of

²² SCHWYZER 1969, p. 456.

Norton's views, one of the many ways that the Schwyzer's thought experiment could be converted into the argumentative form have is as follow:

1. In Ruritania two men are playing chess according of its rules.
2. But in Ruritania chess is not a duel or battle, rather, chess is not a competitive game.
3. If (1) is true, one of the players must win while the other must lose, or occasionally the game ends in a draw (background assumption).
4. But by (2) and (3), if the Ruritania's chess is not a competitive game, then there cannot be a winner and a loser, and the game cannot end in a draw.
5. Therefore, (1) must be false, because chess is a competitive game.
6. Therefore, in Ruritania the two men are not playing chess.

If this conversion is correct, then it seems clear that in Norton's view there is nothing more to draw from this argument, because in trying to evaluate the experiment through the logical-argumentative reconstruction what matters is the way in which the experiment is presented and not its content. In this perspective, the Schwyzer's thought experiment loses its persuasive force, because the discursive force was eliminated. But, it is clear, the important outcomes drawn from this thought experiment go far beyond the conclusion drawn from the premises in the formalized version.

In *Meta-Institutional Concepts: A new Category for Social Ontology* (2014), Giuseppe Lorini offered an accurate analysis of the results of Schwyzer's thought experiment. Lorini defends the very interesting idea that the meta-institutional concepts are conditions of possibility of institutions and institutional facts. The chess rules are socially instituted and, as such, are regulated by concepts that are beyond their own constitutive rules. It is important highlight that in the Lorini's perspective the concept of victory is presupposed by the rules of chess, although it is not constitutive of chess.

According to Lorini's analysis, the activity of the two practices (the sacred rite of chess and game of chess) is governed by the same rules: the constitutive rules of chess.

The difference between the rite of chess and the game of chess lies not at all in what goes on at the board, nor in anything in the minds of those who make the moves. Rather, the difference lies away from the board altogether, in what Wittgenstein calls

the “role” of the activity in the lives of those for whom it is a practice. It is this that makes the one a sacred rite and the other a competitive game. And the difference in roles is shown by what happens, for example, before and after the activity at the chessboard; it is determined by the different ways of speaking and behaving which are appropriate with respect to each case. But the rules of chess can govern what happens only on the chessboard; they can regulate behavior only while the activity is in progress. That is why it will not do to say that “the activity of playing chess is constituted by action in accordance with these rules”.

All the particularities described by Lorini in this passage compose the persuasive force of the thought experiment and its result leads us to an understanding that could never be reached by the “view of the argument”. So, the Schwyzer’s hypothetical scenario is the existence of a community on which the “players” do not seem to consider chess as a competitive game, although they follow the constitutive rules of chess. The evaluation of the hypothetical scenario is that these rules are not enough to determine whether someone really play chess, since meta-institutional concepts such as “victory” or “defeat” are needed²³. Besides this, the Schwyzer’s example also satisfies other conditions which sometimes are required for evaluating thought experiments: the scenario described is fully imaginable (it is coherent and resolvable); the scenario described is applicable (it reveal a characteristic of our conceptual system); and the outcome reveal what Schwyzer wanted to reveal.

This way of evaluating the Schwyzer’s case allows to observe at least two objections that have been raised against certain philosophical thought experiments: (i) objections about unimaginability (that are concerned with the coherence of the scenario described) and (ii) objections about inapplicability (that are concerned with the power of conclusions draw of the hypothetical case and whether they really can be applied in the way that this case was designed).

These challenges are often raised (1) because the scenario cannot be fully imagined, or (2) because it is difficult to know how to make it meaningful, or (3) because we would not be able to apply its lessons to our world²⁴. In this sense, someone might try to object the Schwyzer’s thought experiment because we cannot really imagine one community where chess is not a

²³ The idea of meta-institutional concepts had been advocated by Giuseppe Lorini in many works. Lorini’s perspective has provided another way of thinking about the nature of regulative rules and social ontology. For more detail see: LORINI 2014.

²⁴ GENDLER 2000, pp. 22-3.

competitive game; that although it's possible to imagine such situation, if we try to reason about how we would react to this possibility, we will inevitably make judgments based on assumptions that would not be relevant under such circumstances; although we can imagine that really exists one community for which chess is not a competitive game, and from this we can make informative judgments about the sense that chess have to this community, the conclusions that we can draw based on those judgments do not establish support for meta-institutional concepts. But these objections seem to be linked to the view of the thought experiment as a disguised argument.

Although we have not outlined here a general classification of thought experiments (even because there are already many interesting classifications), we may have a clear idea of their roles in philosophical inquiry. Moreover, it seems possible to show how the imagination can lead us to new knowledge about the world even without the addition of new empirical data. We can learn about our meta-institutional concepts by trying to make sense of hypothetical reasoning presented by Schwyzer, and we can to evaluate actual cases by *thinking* about these hypothetical reasoning. In this sense, the particularities of a hypothetical reasoning can help us to see how the variation of the factors (chess is not a competitive game, for example) could contribute to determine the role that our concepts are playing. If it is correct, then the careful examination of thought experiments could provide us with new knowledge in a way that argument alone cannot.

The new knowledge derived by thought experiments arises of an act of introspection that brings to light true beliefs (like “hinge propositions”) about the world that until then were unarticulated (maybe because for lack a suitable background in which to make sense of them) and that until then were considered implausible. What this perspective claims, against the Norton's theory, is that the analysis and appraisal of a hypothetical reasoning need not involve reconstructing it explicitly as an argument. More important than the validity of the argument that may eventually be derived from a thought experiment, is the validity of the procedure adopted to draw a relevant conclusion from the experiment.

By describing appropriately selected imaginary scenarios, thought experiments provide contexts within which sense can made of previously incomprehensible conceptual distinctions. And by describing specific situations, hypothetical reasoning can justify conclusions about specific cases without explicit or implicit appeal to more general absolute principles. Besides this, thought experiments have an important persuasive force that cannot be obtained by argument alone. This also shows that the main role of thought experiment is to provide “understanding” about the world (in the Schwyzer's

case, an understanding about the limits and scope of the constitutive rules of chess). In the last section of this paper, I will outline some “correction criteria” for constructing thought experiments that do not involve converting them into arguments.

Another interesting strategy to think about the role of thought experiments has been defended by Experimental Philosophy, whose idea is that thought experiments *track* intuitions. Of course, the term “intuition” in general is metaphysically charged, but there is no place here for an examination of the most diverse characterizations that it has acquired over time. For my purposes, I will take “intuitions” as a sort of propositional attitude, which means that “intuitions” are dependent on propositional knowledge and logical reasoning. In other words, intuitions are like “vehicles” of access to our *background knowledge*. In this sense, thought experiments can be useful when we want to understand what people think about a given scenario (or what intuitions they entertain when confronted with an imaginary situation).

In *Normativity and Epistemic Intuitions* (2008), Jonathan M. Weinberg, Shaun Nichols and Stephen P. Stich, use several thought experiments (commonly used in analytical epistemology) for explore attributions of knowledge. I do not want to make any judgment about the results obtained in these experiments, but only draw attention to the fact that thought experiments can be useful to know the intuitions that people entertain about certain seemingly well-established philosophical theses. These authors argue that the epistemologists must offer an answer to cultural diversity in epistemic intuitions. In my opinion, the cultural diversity about attributions of knowledge can be easily explained by the different theories (or world views) that people entertain in different cultures.

Weinberg and his colleagues explore a very interesting theory (advocated by Richard Nisbett and his collaborators) that there are large and systematic differences between East Asians and Westerners on a long list of basic cognitive process, including perception, attention, and memory. Other differences are demonstrated also in the way that these groups describing, predicting, explaining events, categorize objects and revise beliefs²⁵. According to Nisbett and his colleagues, the “holistic thought” predominates among East Asians, while the “analytic thought” is a prevailing pattern among Westerners. A plausible hypothesis is that all these important intercultural differences in epistemic intuitions can be explained by evaluating the “hinge propositions” (to use a term of Wittgenstein) that make up the background knowledge of a community. In this sense, some thought experiments would be able to trace

²⁵ WEINBERG *et al* 2008, p. 32.

such intuitions against the proper background knowledge to justify those differences.

Weinberg and his colleagues used a version of Dretske's thought experiment to show that epistemic intuitions vary from culture to culture (in this case, the experiment was applied to Westerners and Indians):

Mike is a young man visiting the zoo with his son, and when they come to the zebra cage, Mike points to the animal and says: "that's a zebra". Mike is right – it is a zebra. However, as the older people in his community know, there are lots of ways that people can be tricked into believing things that aren't true. Indeed, the older people in the community know that it's possible that zoo authorities could cleverly disguised mules to look just like zebras, and people viewing the animals would not be able to tell the difference. If the animal that Mike called a zebra had really been such a cleverly painted mule, Mike still would have thought that it was a zebra. Does Mike really know that the animal is a zebra, or does he only believe that it is?²⁶

It seems clear that this thought experiment is not merely analytical, as we discussed earlier, because in addition to ending with a question (which makes it difficult to translate the experiment into an argument without the loss its persuasive force) it aims to trace the intuitions that people of different cultures (or even different socioeconomic status in a same culture) may have. The result of the experiment by Weinberg and his colleagues shows that there is a significant difference between the Western and Indian knowledge attributions.

The intuitions that emerge in this thought experiment are not belonging to a metaphysical realm of universals and the relations between them, as some authors have argued²⁷. It seems more plausible to assume that we have hinge propositions that function as background knowledge and that our intuitions that emerge in thought experiments are vehicles to track them. We could say that the hinge propositions would be "activated" or "rescued" when we are engaged in certain thought experiments. According to Brendel, "some intuitions are relatively stable and commonly shared, because we belong to the same biological species and to cultural and scientific communities with

²⁶ WEINBERG *et al* 2008, 22.

²⁷ James Robert Brown defends that intuitions belong a special faculty which function as a vehicle for direct access to universals and laws of nature. The "platonic thought experiments" can offer us a new knowledge that is not based on new empirical evidence nor is it merely logically derived from old data (BROWN 1991, 77).

shared knowledge”²⁸. However, there are intuitions that are neither intercultural nor intraculturally stable, because they also depend on our changing experience and knowledge.

Concluding Remarks

As we have seen so far, there are several ways of devising thought experiments and their correction criteria. I want to emphasize that although not all thought experiments can be converted into arguments without loss of persuasive force, some can be perfectly converted (especially scientific thought experiments). In other words, we can make non-problematic use of our usual rules of inference to evaluate the correctness (or validity) of thought experiments that can be converted into arguments. But it remains to know how to evaluate the other types of thought experiments that cannot be converted into arguments.

The class of thought experiments that does not lend itself to argumentative conversion has been called by Daniel Dennett of “intuition pumps”. According to Dennett, intuition pumps aren’t formal arguments that prove conclusions from premises, but rather, they are like stories or little fables (imaginative reflections) that supposed teach us something. And they can lead us to an intuition where we can reach an understanding act like “Oh yeah, it has to be like that”. Dennett argues that intuition pumps are “persuasion machines” that philosophers have been using for several thousand years²⁹. This description seems useful because it shows us that we need correction criteria for the use of this kind of thought experiments.

According to Brendel, the proper use of thought experiments “involves the study and evaluation of a situation in which data is changed, modified or reorganized in a planned and controlled way with the help of specific background hypotheses or theories”³⁰. In this sense, if we invent a hypothetical scenario in which we manipulate or change data in a strange way, the effects of these variations should always be under control. In other words, we should always keep in mind how these variations can affect other implicit premises or assumptions of the hypothetical reasoning and whether it is still possible to achieve the desired conclusion. If these conditions are not considered, the thought experiment can suffer from Brendel calls “under-determination”, which makes it difficult to draw a conclusion because some

²⁸ BRENDEL 2004, p. 96.

²⁹ DENNETT 1984, p. 12.

³⁰ BRENDEL 2004, p. 99.

relevant aspects are unclear. When that happens, the thought experiment can be turned into an illegitimate “intuition pump”.

There are many examples of successful thought experiments that help us to clear up the nature of our concepts. These thought experiments (like the Gettier’s cases, for example) do not employ fictional or unusual situations without considering the control of the variables. The Gettier’s cases show how thought experiments can be legitimately employed to clarify a crucial philosophical concept (knowledge as justified true belief). Other philosophical thought experiments “where concepts are used in unfamiliar or strange ways we are sometimes uncertain as to whether the application of a concept is legitimate or not”³¹. To avoid illegitimate intuition pumps, it is necessary pay attention to the fact that no relevant aspects of the imaginary scenario can be under-determined. In other words, if we change data in an imaginary scenario, we must to know how they can affect other aspects assumed in the thought experiment. In this sense, we need to be sure that such changes will not affect the desired outcome.

According to Brendel³², although intuitions seem to be immediate and spontaneous (like an “insight”), they are propositional attitudes shaped by our experiences and by our adaptations to the environment. They are fallible and unstable and run the risk of being systematically misleading. In this sense, thought experiments can appeal to intuitions in a way that can make us believe in conclusions that are not justified. This means that we should to employ hypothetical reasoning carefully to avoid illegitimate intuition pumps.

In answer to the question that appears in the title of this work, we can say that not all thought experiments are disguised arguments, but it is possible to establish criteria for the correct application of those thought experiments that cannot be converted into arguments. This kind of thought experiments are not presented as arguments, but rather as invitations to engage in a constructive activity that aims to make exceptional cases meaningful. But I agree that much work still needs to be done in this direction.

References

BISHOP, M. “An Epistemological Role for Thought Experiments”. In: Niall Shanks (ed.), *Idealization IX: Idealization in Contemporary Physics*. Poznan Studies in the Philosophy of Sciences and Humanities 63, 1998, p.19-33.

³¹ BRENDEL 2004, 103.

³² BRENDEL 2004, 107.

BRENDEL, E. "Intuition Pumps and the Proper Use of Thought Experiments". In: *Dialectica*, vol.58, n.1, p.89-108, 2004.

BROWN, J. R. *The Laboratory of the Mind*. London: Routledge, 1991a.

_____. "Thought Experiments: A Platonic Account". In: Horowitz, T/Massey, G.J. (eds.): *Thought and Thought Experiments. Philosophical Studies* 45, 1991b, p.431-444.

DANCY, J. "The role of imaginary cases in ethics". In: *Pacific Philosophical Quarterly* 66, 1985, p.141-153.

DENNETT, D. C. *Elbow Room*. Oxford: Oxford University Press, 1984.

_____. "Intuition pumps". In: John Brockman (ed.), *The Third Culture: Beyond the Scientific Revolution*. New York: Simon & Schuster, 1995, p.181-197.

GENDLER, T.S. *Thought Experiments: On the Powers and Limits of Imaginary Cases*. New York: Garland Publishing, 2000.

GETTIER, E.L. "Is Justified True Belief Knowledge?" In: *Analysis* 23, p.121-123.

GOODMAN, N. *Fact, Fiction and Forecast*. Indianapolis: Bobbs-Merrill, 1965.

KNOBE, J., NICHOLS, S. *Experimental Philosophy*. Oxford: Oxford U. Press, 2008.

KUHN, T. "A Function for Thought Experiments". In: Kuhn, T.: *The Essential Tension*, Chicago: University of Chicago Press, 1977.

LETHONEN, T. "Idealization and Exemplification as Tools of Philosophy". In: *E-Logos Electronic Journal for Philosophy*, 2012.

LORINI, G. "Meta-Institutional Concepts: A new Category for Social Ontology". In: *Rivista di Estetica*, Torino, 2014, p.127-139.

MACH, E. *Über Gedankenexperimente. Zeitschrift für physikalischen und chemischen Unterricht*, 1897, vol. 10, p.1-5.

NISBETT, K. PENG, K. CHOI, I, NORENZAYAN, A. "Culture and Systems of Thought: Holistic vs. Analytic Cognition". In: *Psychological Review*, 2001.

PAPINEAU, D. "The Poverty of Analysis". In: Carmo, J. *A Companion to Naturalism*. Pelotas: Dissertatio Series of Philosophy, 2016.

SEARLE, J. "The Logical Status of Fictional Discourse". In: *New Literary History*, Vol. 6, n.2, 1975, p.319-332.

SORENSEN, R. A. *Thought Experiments*. Oxford: Oxford University Press, 1992.

WILKES, K.V. *Real People: Personal Identity without Thought Experiments*. Clarendon Paperbacks. Oxford: Clarendon Press, 1994.

WILLIAMSON, T. *Philosophy of Philosophy*. Oxford: Blackwell Publishing, 2007.