

Thought Experiments as Social Practice and the Clash of Imaginers

DANIELE MOLINARI*
University of Parma, Parma, Italy

In the last few years, several philosophers have highlighted the social dimension of imagination. In this paper I argue that thought experiments prompt social uses of imaginings if we understand them as props in games of make-believe. In prescribing to imagine stories that develop through fictional narratives, authors of thought experiments prompt their readers to engage in the same imaginative project—at least in its salient aspects—and to endorse their conclusions. Contributions on this topic focus on cases where coordination across imaginers is immediately successful. However, this is not the end of the story. I draw attention to situations where this is not the case, as the practice of thought experimentation often proceeds through criticism, rejections, and amendments. I focus on cases where imaginers do not endorse the conclusion proposed by the author of a thought experiment and either (i) fully reject the principles of generation, (ii) draw different fictional truths from the same principles, or (iii) amend the principles. Although cases of imaginative disharmony are usually dismissed as failures, I acknowledge them as fruitful steps in the cognitive advancement achievable by thought experiments. Cooperative imaginers challenge the rules of the game in meaningful ways, which leads to enhancing fictional scenarios and framing them through different perspectives.

Keywords: Thought experiments; imagination; Kendall Walton; metaphilosophy; epistemology of imagination.

* I wish to thank Wolfgang Huemer, Mike Stuart, Alfredo Vernazzani, Giulia di Rienzo, Dalila Garbellini, Tommaso Piazza and two anonymous referees for their significant comments on early versions of this paper.

If one takes it to be self-evident that people take pleasure in their own imaginations, then one should remember that such imagination is not like a picture or a three-dimensional model, but a complicated pattern of heterogeneous components: words and images. [Once one does so] one will then no longer oppose operating with written or acoustic signs to operating with “mental images” of events. (Wittgenstein 2018: 44)

1. *Introduction*

We usually think about imagination as a private, creative and unconstrained mental activity. And when we do so, we often have in mind examples of a similar kind, such as: the daydreamer who imagines herself drinking a refreshing cocktail on a Caribbean beach; the artist who mentally explores her next steps in creating an inspired artwork; the child dreaming a fantastic kingdom where the birds fly in the underground. All these imaginers are engaged in silent, lonesome imaginings. Nevertheless, if we consider other cases, we can focus on imaginative acts through different lenses. Take for instance the children playing cops and robbers in the garden, the actors performing in improvisational theatre, role-playing games and some kinds of shared meditation. These cases of joint activities seem to require participants to take part in the same imaginative projects in order to be successful. Thus, what strikes as salient in a heterogeneous family of activities such as imaginings (Kind 2013; Murphy 2020a) depends on which examples we take into consideration.

In recent literature, more attention has been paid to social aspects of imaginings in general (Walton 1990; Szanto 2017), in architectural practices (Murphy 2004, 2005), in scientific models (Salis 2020; Salis and Frigg 2020) and in thought experiments (Meynell 2014, 2018; Bancog and Song 2020; Salis and Frigg 2020) among others. In this paper I will follow this path and focus on the social dimension of thought experiments. Most of these accounts are based on Kendall Walton’s groundbreaking *Mimesis as Make-Believe* (1990). Even though they all acknowledge the relevance of the social aspects of imagination in thought experiments, there are several issues that have not yet been properly addressed and wait for further clarification. In what follows, I will build my argument on the Walton’s theory as well, but highlight an aspect that is neglected in the actual literature: taking thought experiments as a case study of social imaginings can shed light on the dimensions of clash and disharmony in imaginative projects, even in collaborative ones. As I understand it, social imagining does not presuppose harmony; we can—and we often do—imagine together with others even when there is disagreement between imaginers. This point will turn out to be an epistemic virtue, as divergencies in the conclusions of thought experiments can help in refining the fictional scenario and the issue at stake.

Thought experiments are quite compelling and seductive insofar as they unfold through rhetorical ornaments and fictional elements. Some of them are even so popular that they can be regarded as pop culture stuff, such as Schrödinger's Cat. However, the authors' ability to illustrate their point using this kind of device can lead to some suspicion. In the history of philosophy, we can find several concerns about the possible misuses of fiction and imagination in the understanding of reality; the power of *pictures* to "hold us captive" (Wittgenstein 2009: §115) is well known, after all. An appealing thought experiment may, therefore, prompt us to endorse a well-written, but flawed conclusion.

In this paper I suggest that this concern is somewhat overrated, as the social practice of thought experiments often encourages researchers to criticize and challenge an author's conclusion. Thought experiments are not only successful devices for illustrating or arguing for a thesis (among others uses); they are also *dialectical moves* that allow even those who do not share their conclusions to take a step forward and enrich the debate. The cognitive value of thought experiments also lies in their prompting of criticism, insofar as the clash they encourage is epistemically productive.

The paper develops two central arguments: that (i) thought experiments prompt social uses of imaginings if we understand them as props in games of make-believe and that (ii) cases of imaginative disharmony are at the heart of thought experiments as social practice and, thus, should be considered fruitful steps in cognitive advancement. The two arguments are intertwined by the topic of rule-following, which will stay in the background as an underground river flowing throughout the paper. The first argument will show that imagining together is (also) a matter of complying with the rules set by objects designed for this task by their authors. The second one, on the other hand, focuses on the capacity of the participants in the imaginative activity to break these rules, and will suggest that this *anarchic* activity can lead to an epistemic progress.

The paper is structured as follows: section 2 is dedicated to the Waltonian theory in order to highlight the social and normative aspects of our imaginative activities, along with the objects involved in them. In section 3 I will focus on thought experiments within this framework, understanding them as props with the social function of prescribing meaningful imaginings. Section 4 will be the core of my paper. Here I will consider the different ways in which a thought experiment can be criticized, emphasizing the epistemic value of the clash between researchers engaged in the same thought experiment.

2. *The social dimension in games of make-believe*

According to Walton (1990), games of make-believe are imaginative activities in which we explore fictional worlds. For example, some children who play together imagine that the floor in the living room is deadly

lava. Someone among them shouts “Watch out! The floor is lava!” and the fictional adventure quickly sets in: the children immediately jump on chairs and sofas to protect themselves.¹ From the beginning of the game, and until its conclusion, the children imagine that “the floor is lava” and behave accordingly.² To make-believe that a given proposition is true (that it is a *fictional truth*) implies some sort of imaginative constraints to it, in more or less rigorous ways depending on the kind of game being played. What is true in a fictional world is constrained by the rules that all players must accept in order to participate in the same imaginative activity. Walton calls these rules *principles of generation* (Walton 1990: 38). Thus, if the children accept the principle of generation “the floor is lava”, they begin to share the same fictional world—that is, they all start a game in which it is prescribed to imagine that the floor in the living room is made of lava.³ As long as they keep playing together, their actions and imaginings are constrained by the principle(s) of generation that they have mutually agreed upon. Accordingly, the action of a child walking on the floor can have several meanings within the game. For example, she either (i) imagines to be tired of living, (ii) proposes a new principle of generation and pretends to wear shoes with special soles, (iii) does not play correctly or (iv) is just bored and decides to quit the game.

We can thus highlight a salient feature in games of make-believe: they introduce “criteria of correctness” in imaginative activities. No player is allowed to imagine everything that just pops into her head during make-believe. There are rules that determine what is true in the game (i.e., *fictional truths*) and authorize certain kinds of imaginings and not others.

Far from being exclusive to children’s games of make-believe, the principles of generation are also central in all forms of representational art, like sculptures, paintings, movies and novels among others. All

¹ This is just one of many possible ways to start a game of make-believe. For example, a child might say nothing, but act as if she was in pain on contact with the floor, saying something like “gee, that was close!”. Depending on the reaction of the other children, the game will either begin or be rejected. Thus, it is not necessary to explicitly formulate any proposition to start a game of make-believe.

² The children in this example do not only imagine propositions: they are imagining that “the floor is lava” as well as imagining objects (such as lava) and actions (such as saving their own lives). These three kinds of imagination (called “propositional”, “objectual” and “experiential” imagination, respectively) can all be prompted in games of make-believe (Walton 1990: 42 f.).

³ Unlike possible worlds *à la* Lewis (1986) in which every proposition has a truth value, fictional worlds are indeterminate in many aspects and may contain contradictions and other absurdities. Some fictional worlds even prompt our imaginings because of their indeterminacy. Many literary minimalist stories, such as those written by Raymond Carver, prompt meaningful imaginings precisely because they remain silent about the consequences of certain fictional actions. But what are fictional worlds? Walton seems to conceive them as collections or clusters of fictional truths, although this does not imply any ontological commitments.

these works of art prescribe different kinds of imagining (to their appreciators). For example, Maurits C. Escher's *Relativity* prescribes the viewer to imagine a tangled stairwell, while the incipit of Raymond Carver's short story *After the Denim* asks the reader to imagine that "Edith Packer had the tape cassette plugged into her ear, and she was smoking one of his cigarettes" (Carver 1989: 67).

According to the examples presented so far, we obtain and use principles of generation through a wide variety of objects and levels of sophistication; from improvised games to artistic masterpieces, we can recognize some kind of relationship between principles of generation and concrete objects. In the context of the first example, if children were in the backyard, the principle "the living room floor is lava" would sound wrong or unworkable. It seems to assign the concrete living room floor an essential role in the imaginative project. According to Walton, principles of generation generate props, that is, objects which, in turn, generate fictional truths; they determine what is true in the world of fiction. The principles of generation, thus, are prescriptions to imagine, which the participants of the game need to comply with. Props are objects that retrieve principles of generation and that can give coherence to a fictional world. For the children who participate in the game, the floor becomes a prop, just like Escher's lithography and the copy of Carver's book become a prop for the audience. Employing props in imaginative activities makes it possible to ground the games of make-believe on objects that can be intersubjectively perceived and enjoyed by all the participants in the game. Floor tiles, prints and texts may serve as external criterion in this.

Unlike floor tiles, an object specifically designed for being used as prop—such as a painting or a fictional narrative—is always associated with its own fictional world (or its own cluster of fictional truths) that participants of the game are invited to imagine. Even if someone, engaging in the game prompted by Escher's lithography, would imagine go-kart tracks instead of chaotic stairs, the world of *Relativity* would not change at all—that is, the imaginer's failure to conform her imagining to the prescriptions contained in the work does not change the fictional world generated by it.

In this context it may be helpful to consider the distinction that Walton makes between the "work world" and the "game world". The player builds her own world of fiction by importing and expanding that of the work. Leaving extreme examples aside (imagining go-kart tracks albeit the prop asks us to imagine stairs can be understood as a refusal to cooperate), even in most accurate cases it is possible to find some minimal discrepancies between the work world and the game world that, however, do not undermine the quality of the game and the players' coordination.⁴ The cluster of fictional truths associated with a work

⁴ For example, in a game world it could be true that the viewer observes a couple arm-in-arm walking on the stairs of Escher's *Relativity*, while the same proposition

world is an external and objective criterion by which imaginers can “calibrate” their imaginings. But not all fictional truths are explicit and straightforward during the contemplation of a fictional world. In this regard, game worlds can be useful as their “authors” (the players in the game of make-believe) can sometimes focus on some fictional truths rather than others, grasping something meaningful and offering it to other players—at least to cooperative imaginers.⁵ Work worlds are, therefore, “out there” and ready to be explored.⁶

The salient point here is that the work world has a normative aspect that all players should take into account in order to properly play the same game. On the other hand, the game worlds are a plurality of worlds—one for each player—which may conflict with each other, and in which personal peculiarities can be put into play. The work world is, therefore, the world of conformity to rules, whereas game worlds are more anarchic. The work world sets the constraints and, depending on the purpose of the game of make-believe, participants can be encouraged to challenge these constraints in meaningful ways within their own game worlds.

Nonetheless, it can sometimes be difficult for all participants to take part in the same game. The meaning and the application of principles of generation may indeed differ, depending on the target community. It may be difficult for a New York broker and a North Sentinel inhabitant to play the same game but, after all, they would have more basic difficulties to communicate in the first place.⁷ However, problems can also arise between players with similar cultural backgrounds. Here we can find more widespread and interesting cases of partial incomprehension, that is, when one successfully participates in the game but imagines different details or draws different implications. If a narrative prescribes to visually imagine two falling objects tied to one another without further instructions, then there are many aspects that remain blank and can be filled in by the reader by her own will or her

is not true in the work world. The viewer is not inside the lithograph observing people, so “the viewer observes a couple arm-in-arm” is not a fictional truth in the world of this artwork. This difference between a game world and the work world does not spoil the prop insofar as the viewer is able and willing to follow its prescriptions.

⁵ In this context, a cooperative player is whoever intends to follow the work world in creating her own game world and is open to suggestions from other players who intend to follow the work world as well.

⁶ This means that work worlds are independent from their authors. Even the author herself, playing the game of make-believe prompted by her creation, constitutes her own game world. In other words, who designs a prop is just another player with no specific privileges.

⁷ I assume that if communication is at risk, then it is difficult to start joint activities based on imaginings. After all, the whole make-believe mechanism is based on the ability to prescribe imaginings. If the props are not apt to convey these prescriptions, and the player is equally unable to retrieve and understand them, then it would be difficult to start any collaborative imaginative project.

idiosyncratic preferences. I can imagine spiky blue rocks, tied by a jute rope, while another player can imagine gray smooth spheres, tied by a cotton rope. We would both correctly follow the same prescriptions. To what extent one can freely imagine the details is determined by the aptness of a prop's prescriptions.

Games in which we have to imagine down to the smallest detail require equally accurate prescriptions—and, thus, excellent authors. If a narrative prescribes us to imagine two falling bodies in order to argue against Aristotelian physics, then it will not add to the aesthetic details of the bodies but highlight other qualities, such as their different sizes and their being composed of the same material. In this case, as long as there is consensus on which principles of generation apply, differences in game worlds based on personal idiosyncrasies do not undermine the participants' imaginative harmony. In other words, the participants in the same game of make-believe will imagine the same sequence of events, at least in the aspects the narrative makes salient. If, on the other hand, the prop is designed to solicit a relatively detailed visual mental image, then complex descriptions or other media (such as paintings) will be employed to prescribe the most suitable imaginings.

Even when participants' game worlds diverge on relevant details, however, such dissonance may not be a problem. Some props (such as photographs) may be apt to prescribe *de se* imaginings (imaginings about oneself) with the aim of prompting one's memories, and this would trigger different streams of imaginings in each player. A certain amount of ambiguity might be desirable, as well as deliberately pursued by the author of the prop, as it might be significant in some respects. Indeterminacy and sketchy scenarios can encourage the creation of insightful game worlds. Their authors can then communicate the peculiarities of their game worlds to other players, suggesting implications or highlighting aesthetic details that others could have overlooked (Meynell 2018: 504). The lack of harmony between imaginers may be a matter of time. Moreover, it could be useful to discuss which rules to accept or which consequences to draw from them. In this case, ambiguity would be an effective rhetorical device in prompting the creation of new principles of generation, the quality and relevance of which would rely in the players' imaginative capacities. In section 4 of this paper I will return on the disharmony between imaginers, focusing on its potential epistemic value.

Let me draw some conclusions from what has been written until now. Walton's theory fruitfully highlights a marked normative and social dimension of our imaginings. It does so by focusing on concrete objects involved in imaginative projects and showing that props can coordinate our imaginings, giving participants a sense of the rules and constraints at play in the imaginative activity. Emphasizing the normative and social level of imagination allows Walton to draw an analogy between shared imaginings and games, shifting the attention from

our personal and lonesome imaginings to activities guided by intersubjectively valid criteria. Whoever does not follow the rules is free to do so but is out of the game.

So far, we have appreciated the social aspect of imagination. However, the activity of imagining together can be investigated in different ways. An alternative to Walton's theory is Keith Murphy's notion of "collaborative imaginings" (Murphy 2005). The anthropologist, like Walton, acknowledges a crucial role to concrete objects during shared imaginings. He invites us to consider the example of a group of architects who design a service yard section of a laboratory building. With this example, he highlights the central role of concrete objects in shared imaginings insofar as participants work on the same project employing drawings, gestures and verbal suggestions. According to Murphy, the building draft map placed on the work table "serves as the actual anchor of the talk" (Murphy 2005: 124). Moreover, the architect who places his hand on the map, exactly where he would like to locate a door, is using a gesture in order to help other participants to collectively imagine how the building would be modified according to his suggestion.

However, a make-believe-oriented account is more inclusive than the one provided by Murphy, as the latter considers perceptual-like imaginings only. According to Murphy, collaborative imagining is a special kind of perception—or something like a bridge between visual perception and imagination—which he calls "*perceiving in the hypothetical mode*, that is, purposefully seeing things as if they were something else" (Murphy 2005: 117). This characterization of imagining can easily account for the specific architectural activities under consideration, but it does not account as well for other kinds of shared imaginings such as those based on fictional narratives. They can call for different kinds of imagination depending on which prescriptions are at play. A detailed description of a castle on the clouds, for instance, could prompt perception-like imaginings resembling René Magritte's *Castle in the Pyrenees* while an intimate story narrated from an engaged point of view could trigger forms of experiential imaginings.

Moreover, Murphy's account succeeds in explaining a synchronous kind of imaginative activities in which participants perform collective imaginings at the same time. The simultaneity of participation in these architectural activities encourage the dialogue and, therefore, prompts a coordinated and engaged discussion based on shared imaginings in synchrony. However, we can imagine together also in asynchrony, such as when we discuss the same thought experiment, criticizing it or designing some of its variants, at different times. Walton's theory has the merit of successfully explaining these cases as well, as props keep their capacity to prescribe imaginings to different audiences at different times. Asynchronous imagining together can easily explain how the debate on a thought experiment (or other fictional narratives) actually

works, that is, through dialectical moves that occur in different temporal stages.⁸

The distinction between synchronous and asynchronous kinds of imagining together can correspond to the two meanings of “together” which Thomas Szanto distinguishes in “imagining together”. It is one thing when some researchers non-simultaneously imagine something, another when they imagine the same state of affairs at the same time. The former is called “imagining something *alongside* with others” while the latter is a form of “*collectively* imagining something *together*” (Szanto 2017: 232). Szanto, as well as Murphy, focuses on the second case—undoubtedly more interesting on a phenomenological level of analysis—and claims that the first one is just an “ordinary case” of social imaginings. However, my line of reasoning proceeds on a different level. I do not focus on imagination as a mental act, neither private nor collective, but on its uses within philosophical and scientific research communities; and one of the most widely used imagination-based devices employed in these contexts are thought experiments. It is for this very reason that I will delve into the topic of thought experiments in the next section.

3. *Thought experiments as social practice*

What thought experiments are is a matter of debate. Very briefly, they can be understood as arguments “disguised in a vivid pictorial or narrative form” (Norton 2004: 45), as “telescopes into the abstract realm” of Platonic entities (Brown 2004: 1131), as mental models that reconfigure past experiences through simulation and memory (Gendler 2010; Mišćević 1992, 2007; Nersessian 1992, 2018), or as a special kind of fictional narratives (Carroll 2002; Meynell 2014, 2018; Salis and Frigg 2020; Willée 2019) among others. In this paper I will focus on the fictionalist account, since understanding thought experiments as fictional narratives highlights the essential role that imagination plays in them.⁹ Walton’s theory, unlike other ones, allows me to acknowledge

⁸ An interesting feature of the asynchronous kind of imagining together is that it enables an imaginative project to proceed even after its author’s death.

⁹ Most fictionalist accounts compare thought experiments and literary works of fiction. This analogy is at the heart of an entire research program, and it has been fruitfully discussed by several philosophers. For instance, Catherine Elgin focuses on the mechanisms of exemplification at work in both kinds of artifact (2014). David Egan, on the other hand, argues for a skeptical outcome, insofar as—unlike literary works—thought experiments are always used to make arguments (2016). Moreover, Iris Vidmar highlights the cognitive value of hypotheses that can be found in both thought experiments and literary fiction (2013) and, more recently, Alice Murphy concentrates on the aesthetic details used in thought experiments and the flexibility of its interpretations (2020b). Finally, David Davies (2007) provides a conceptual geography in which the main accounts of thought experiments are linked to some central questions in the philosophy of literature. See his (Davies 2018) for an excellent overview of the debate. I choose not to elaborate on this analogy in my

the social dimension of thought experiments, and to explain how researchers engaged in thought experiments can take the same fictional event into consideration: they participate in the same game of make-believe—or, in other words, they comply with the same prescriptions to imagine.¹⁰ The section of John Searle's paper in which the *Chinese Room* thought experiment is presented (1980: 417 f.), for example, is a prop that generates, among others, the fictional truth that the person locked inside the room can answer questions formulated in Chinese even though she does not speak Chinese.¹¹ Similarly, in Derek Parfit's *Reasons and Persons* there is a prop that invites the reader to imagine a science fiction scenario in which teletransportation is real (Parfit 1984: 199 f.).¹²

Both props are specifically designed to claim that computers operate syntactically with no semantic understanding and that personal

paper, as I have explored elsewhere the similarities and differences between thought experiments and other fictional works (see Molinari 2020).

¹⁰ One might argue that thought experiments could also be conducted by its authors in isolation. However, props will be used in these cases as well: just as a painting could be drawn and kept secret by its painter, an author of a thought experiment could write a fictional narrative and keep it for herself. Actually, in both cases a fictional world will be created. The main difference between these lonesome cases and the social ones is that, in the former, the solitary participant needs only her consent to start, modify, or end a game of make-believe, while in the latter the consent of all players is required. I chose to delve into the social uses of thought experiments because, I argue, this is their primary—as well as their most fruitful—use in philosophical and scientific research communities.

¹¹ Very briefly, Searle asks us to imagine that there is a machine capable of properly answering all questions in Chinese and to pass the *Turing test*. According to strong A.I. theorists, this computer would understand Chinese because of the similarity between its behavior and that of a native Chinese speaker. Now imagine yourself locked inside this machine. In front of you there is a book containing the English version of the program used by the computer, along with plenty of paper and pens. Your task is to receive and send back Chinese ideograms from two openings connected to the outside world. You cannot understand these ideograms. However, thanks to the instructions in the book, you can create new Chinese ideograms as output. In this fictional scenario you're able to create answers that a Chinese speaker would find satisfactory, although you don't understand any of them. All you do is to follow instructions in the book. According to Searle, this lack of understanding suggests that a computer, being in the same situation as you, does not understand anything either.

¹² Parfit's famous *Teletransporter's* thought experiment invites us to imagine a futuristic scenario in which teletransportation has been invented. This technology copies all the cells in a person's body, immediately sending them to another planet. An exact copy of the body is created at the arrival station. During this operation, the original one is destroyed. Parfit elaborates this fictional scenario wondering what it can show us about the concept of personal identity. There seems to be no problems at this stage of the story: if you enter the teletransporter, you're still yourself but on Mars. However, Parfit modifies the unfolding of events, making it difficult to know which person to identify with. Imagine that, due to a technical problem, your original body is not destroyed but is going to die. While your replica has arrived at its destination, your original self is dying. It is a fictional truth either that you will survive in your replica or that you will exist as two people for a few moments.

identity is not an “all or nothing” matter, respectively.¹³ Thought experiments that occur in philosophical or scientific papers are, thus, *fictional narratives* that prescribe meaningful imaginings for epistemic purposes (to claim for, evaluate, show or reject a thesis). In order to be able to perform this function, the reader needs to accept the principles of generation at work, explore what would happen in the fictional scenario and ponders whether what is true in the fiction can also be significant for worldly-cognitive purposes. If the reader is able to figure out which consequences and implications are obtained by the principles of generation, it is possible that she will take into consideration a proposition she had never thought of, or to observe a phenomenon from a new perspective. If this consideration is correct, it shows that the most interesting thought experiments may be understood as some sort of epistemic calls to action—or, in our case, *calls to imagining*.

It is worth noting that Walton’s theory and the mental model account are compatible, though. Among the many types of imaginings that a thought experiment can prompt, there may also be mental modeling—in which spatial or kinetic elements strike as salient. For instance, Hume’s *Missing shade of blue* (Hume 1999: 9 f.) triggers a perceptual kind of imagination, while wondering to imagine whether the sofa can get through the door would prompt the creation of a mental model. Moreover, Nersessian (2018: 313) herself draws attention to the narrative presentation of thought experiments, although she focuses on mental representations it triggers rather than its social and normative dimension.

4. *The cooperative clash of imaginers*

I have highlighted two main points so far. The first is that people can imagine together by using props; the second is that thought experiments, understood as a special kind of props, invite readers to collaborate in epistemic, imaginative projects. Contributions on this topic typically focus on cases where coordination across imaginers is successful at once, that is, when readers accept all the principles of generation designed by the author, play along with the narrative, and endorse the conclusion. However, this is not the end of the story, as the practice of thought experiments often proceeds through criticism, rejections and amendments. Endorsing the author’s conclusion is only one of the possible outcomes: once published, thought experiments (as well as other

¹³ It might seem that there is a gap to be bridged between “concrete props”, such as the aforementioned floor, Escher’s lithograph and Carver’s book on the one hand, and thought experiments on the other. After all, unlike thought experiments, lithographs and books can be hung on a wall or placed on a table. However, understanding thought experiments as fictional narratives can help to bridge the gap. Just like other fictional narratives, thought experiments are props composed of texts, utterances, or objects made up of both texts and images. See (Meynell 2018) for an insightful discussion on thought experiments and the pictures that often accompany them.

props) begin to be conducted and discussed within the community of researchers. These dynamics prompt participants not only to follow prescriptions to imagine, but even to both acknowledge and evaluate the rules and constraints at play in each thought experiment. It is precisely at this point that the social dimension of thought experiments comes into focus and shows an epistemic value—by discussing, modifying or rejecting the imagined scenarios, along with the rules that design them.

The point can be highlighted by analyzing the common practice of criticizing a thought experiment. By framing it within a Walton-inspired theory, criticisms can be done in at least three ways:

- (i) *Presenting new principles of generation.* A critic might reject all the principles of generation and fictional truths made explicit by the author. That is, the player thinks the game is just wrong or misleading. This can happen for different reasons. For instance, the opponent may find that, after accepting these principles and these fictional truths, the proposed scenario is overly implausible, clueless or not apt. In these cases, researchers usually proceed by presenting a new thought experiment that is claimed to be better suited for the epistemic purposes of the discussion. For instance, Searle quotes an objection to his *Chinese room* in which the reader is asked to imagine a new fictional world by presenting a new prop: she is no longer asked to imagine a person locked up in an isolated room but “a program [...] that simulates the actual sequence of neuron firing at the synapses of the brain of a native Chinese speaker when he understands stories in Chinese and gives answer to them” (Searle 1980: 420). The opponent believes, perhaps on the back of previous theoretical commitments concerning strong A.I., that this new fictional scenario is better suited to shed light on the problem, insofar as “at the level of the synapses, what would or could be different about the program of the computer and the program of the Chinese brain?” (Searle 1980: 420). Searle replies to this objection by accepting the new game but framing it in his own theoretical framework and criticizing its conclusion.
- (ii) *Reorganizing fictional truths.* A critic might accept the fictional world but perform a kind of “semantic reorganization” of the elements that were already present in order to highlight other fictional truths. Thought experiments are fictional narratives, and each narrative unfolds through choices: the author chooses what to focus on and what to neglect and, in doing so, she marks certain aspects as salient while hiding others. Thus, a clash between imaginers may arise about which elements are significant and how they should interact in the imaginative project. For example, Searle discusses a second objection in which a different conclusion to his *Chinese room* story is proposed: “while it is true that the individual person who is locked in the room does not understand the story, the fact is that

he is merely part of a whole system, and the system does understand the story” (Searle 1980: 419). Here it is argued that, by accepting all the principles of generation designed by Searle, the prop prescribes the reader to imagine a new fictional truth, namely that all the objects in the room operate as a system. Therefore, Searle’s point needs to be revised. In particular, the philosopher is accused of having made the mistake of focusing his narrative too much on the single person locked up and not on the relations between the person and the other objects. This technique bears resemblances to what Roy Sorensen calls “smartfounding” (Sorensen 2019): the critic does not play along, but resists the thought experiment by showing that, even if one’s imagining does conform to the prescriptions made by the author, one could obtain unexpected fictional truths that cause the imaginative project to fail. For example, the smartfounder engaging in Einstein’s *Chasing the light* thought experiment would not grasp any Einstein’s insight; she would only imagine an instant death caused by travelling at the speed of light (Norton 2013: 123).¹⁴ Actually, smartfounding is not like performing a “semantic reorganization” in search of other significant fictional truths—it is just a way of refusing to cooperate. “Many hypotheticals have minor flaws that cooperative hearers ignore” (Sorensen 2019: 792). Thought experiments’ narratives have such flaws as well, and cooperative imaginers know, following context-specific and epistemic constraints, which new fictional truths turn out to be meaningful and which ones only muddy the waters in the debate.

(iii) *Amending principles of generation.* A critic might both accept some principles of generation and amend others, in the conviction that a similar but not identical fictional world is more insightful for a given epistemic purpose. The technique consists of modifying certain details in order to significantly diverge the unfolding of fictional events and in consequence alter the conclusions. For instance, a detractor of Thomson’s *Dying violinist* (1971: 48 f.) may modify the prop, asking to imagine a sick beloved one instead of a stranger violinist.¹⁵ This detail would introduce an affective bond—absent in the original fictional narrative—that could change the result of the

¹⁴ This tragic and brutal epilogue that results from a particular reading of Einstein’s thought experiment is discussed by Michael Stuart in (2020: 974).

¹⁵ In this famous thought experiment the philosopher asks you to imagine yourself waking up in bed next to a famous violinist who, as you learn right after your awakening, suffers from a kidney disease and risks dying. *The Society of Music Lovers* has kidnapped you because you have the same rare blood-type as the violinist and could, with your circulatory system pumping blood also through the violinist’s body, save the life of the violinist. The hospital director concisely states: to save their life, you have to stay connected to their body for nine months. At this point, Judith Thomson asks the reader: “is it morally incumbent on you to accede to this situation?” (Thomson 1971: 49). This thought experiment invites the reader to imagine a fictional world designed to conceive, by analogy, the possible relationship

thought experiment. We can find another example for this strategy in the replies to Searle's *Chinese room*, in which the isolated room is modified into a robot that, while it keeps receiving and delivering Chinese symbols, behaves indistinguishable from a system that actually perceives the nearby environment. It is worth noting that Searle responds to the objection by amending the fictional world a second time and reintroducing the human being, no longer locked in a room but within the giant robot.

All three techniques—presentation of new principles of generation, reorganization of fictional truths and amendment of some principles of generation—are guided by one's theoretical assumptions or epistemic *desiderata* on the topic at stake, and even by elaborate perspectives or worldviews. They all are strategies to change the spin of a thought experiment or to show its conclusion is false. Moreover, critics do try to build on it to bring home their own message—which shows that they recognize the epistemic value of the thought experiment they aim to criticize.

As it might be clear from the examples above, the whole section dedicated to the replies in Searle's paper is an interesting case study on the social dimension of imaginings. It suggests that a lively discussion has been carried out, in which different fictional worlds were to be imagined (we are asked to imagine systems composed by human beings, books, sheets of paper and pencils; robots equipped with cameras, arms and legs; sci-fi technologies that stimulate neuronal activity among others); in which the participants were required to play along with fictional worlds proposed by the others and in which all the three techniques of critique were applied. Moreover, as a debate concerning strong artificial intelligence, it unfolds across multiple disciplines. It shows, therefore, that sharing imaginings for epistemic purposes is not an exclusive philosophical matter but can occur in scientific contexts as well.

Actually, it should be noted that the point generalizes to most thought experiments that have been discussed within its research community. Putnam's *Twin Earth* thought experiment, for instance, prompts an insightful debate based on the same imaginative project as well. However, the debate it prompts has to be traced between different publications. Searle's paper, on the other hand, is a more accessible example—it is the author himself who quotes and evaluates the criticism directed at his thought experiment.

Nonetheless, all these rejecting and modifying fictional worlds seem to reveal a significant point: thought experiments prompt clashes, as well as harmony, between imaginers. Luckily enough, we can imagine together despite (or rather, thanks to) the underlying disharmony. To highlight this point is precisely the contribution I want to make with this paper.

between a mother and her fetus, and to understand some moral implications of abortion that could easily be underestimated or neglected.

I suggest that these cases of imaginative disharmony, usually dismissed as failures, are at the core of thought experiments as epistemic practice. As we have seen, imaginings in thought experiments (or, more generally, imaginings with others) are also a matter of rule following. That is, the readers engage in the same thought experiment as long as they follow the prop's prescriptions. However, this is just the beginning of the process: researchers who are willing to collaborate on the same imaginative project often do not immediately endorse the conclusion presented by the author of the thought experiment. On the contrary, they test it, evaluate other fictional worlds, break the rules and suggest new ones. In other words, they keep changing the prescriptions to imagine, giving their peculiar game worlds a prominent role. This activity is epistemically fruitful insofar as it prompts the exploration of novel fictional scenarios, the unveiling of some narrative implications neglected by the author, or the acknowledgement of which elements of the story are essential to reach a certain conclusion. As we have seen through the distinction between work world and game worlds, all the participants in the game may come to imagine new meaningful implications—or other significant but overlooked details—by developing their own game world. These dynamics call for the refinement of fictional narratives, the evaluation of alternative analogies and perspectives concerning the issue at stake. If a thought experiment is compelling, then it prompts both its critics to challenge it and its supporters to refine it.

This shows that thought experiments are part of the usual dialectical processes we know from philosophy and the sciences: also arguments do not convince right away (even if they are valid), but trigger forms of response that further elaborate or show wrong a given point: they are but one move in a more complex game. Thought experiments are just *dialectical moves* played in a somewhat different game than that of arguments. Being a different game, it calls for different rules. We already comply with some well-known rules for the advancement of debates through arguments: for instance, if the inferences are valid and the premises are true, then the conclusions will be true as well. Critics of arguments, therefore, will focus on uncovering invalid inferences or false premises.

Critics of thought experiments, on the other hand, do not seem to have such clearly delineated rules to comply with. Yet, thought experiments can be found in almost the entire history of philosophy and are still successfully employed nowadays. Moreover, if I am right, their success is enhanced insofar as they encourage their critics to participate in the same imaginative project and to propose the most various prescriptions to imagine.

The three techniques of criticizing a thought experiment that I have listed above are only little moves in a more complex project, which consists in the theorization of a *deontology* of thought experiments. The

purpose of this normative theory would be to make the rules explicit for successfully constraining our imaginings to conduct the dialectical game of thought experiments, along with its meaningful exceptions.

However, the project of a deontology composed of rules engraved in stone that are valid for all thought experiments could easily turn out to be a philosophical chimera. In a recent paper (2020) Michael Stuart argues that the epistemic power of imagination in the sciences is also to be found in its being *productively anarchic* and that researchers may achieve cognitive advancements precisely by breaking the rules that they (or others) have set for themselves. Nonetheless, identifying a set of malleable and context-dependent rules—even though they do not provide us with a theory that fits all cases—could help us to understand the success of the most famous thought experiments and, perhaps, even some of those yet to come.

The clash dimension intrinsic to the practice of thought experiments should not be understood in terms of resistance to them. Sorensen's excellent discussion on this latter topic in his (2019) highlights different ways of non-cooperation in the very practice of thought experiments, from unschooled response to imaginative resistance and sophisticated smartfounding. In these cases, those who engage in the thought experiment actually refuse the invitation to participate in the game of make-believe—expressing their intention not to cooperate. The causes of non-cooperation can be various. From the inability to recognize any epistemic value in imagination, to a general distrust in it; from the imaginative resistance that can be prompted by having to imagine alienating or unpleasant states of affairs, to the simple desire to make the opponent look like a fool.

My paper, however, focuses on the disharmony of imaginings between cooperating imaginers. That is, between those who have accepted—and not resisted—thought experiments. In order for this epistemic practice to flourish, cooperative researchers design exceptions to the rules of the game, breaking them in the most constructive, meaningful, and even anarchic ways. My point could be seen in continuity with Stuart's aforementioned idea of imagination as—at least in part—productively anarchic, although reframed in the context of social imaginings and collaborative imaginers. If we recognize the epistemic value of the productive anarchy of imagination, then we must take seriously into account the imaginative efforts with which cooperative imaginers challenge the rules of the game, along with their insightful game worlds—in order to evaluate their novel proposals within the research community.

5. Conclusion

In this paper I explored the mechanisms of cooperative clash between participants in the same thought experiment, focusing on their potential epistemic value. To achieve this, I started by pointing out some

important features of the Waltonian theory. Here we came across some technical terms such as “principle of generation”, “prop” and “work world/game world” in order to emphasize the fact that, respectively, (i) we comply with rules to conduct social imaginative games, that (ii) we employ objects to understand which rules are at work, and that (iii) there may be a certain productive tension between the rules set by the objects and the imaginative freedom of the participants.

Framing thought experiments through this theory served two main purposes. Firstly, to emphasize the essential role that imagination plays in these epistemic devices. The second one is to understand thought experiments as dialectic moves in a game of imagination in which not only the author, but also other contributors participate—a game that develops through challenges, criticism and manipulation of fictional narratives.

The three techniques of critique that I have outlined are nothing more than an attempt to explicit some general ways in which a thought experiment can be challenged by a cooperative imaginer. They succeed in showing the importance of what I have briefly mentioned as a *deontology* of thought experiments, that is, they call for further investigation into which ways of regulating our imaginings are appropriate—and in which ways it is appropriate to break the rules instead—when it comes to employ them in our epistemic endeavors.

References

- Bancong, H. and Song, J. 2020. “Exploring How Students Construct Collaborative Thought Experiments During Physics Problem-Solving Activities.” *Science and Education* 29: 617–645.
- Brown, J. R. 2004. “Peeking into Plato’s Heaven.” *Philosophy of Science* 71: 1126–1138.
- Carroll, N. 2002. “The Wheel of Virtue: Art, Literature, and Moral Knowledge.” *The Journal of Aesthetics and Art Criticism* 60 (1): 3–26.
- Carver, R. 1989. *What We Talk About When We Talk About Love*. London: Vintage Books.
- Davies, D. 2007. “Thought Experiments and Fictional Narratives.” *Croatian Journal of Philosophy* 7: 29–45.
- Davies, D. 2018. “Art and Thought Experiments.” In M. T. Stuart, Y. Fehige and J. R. Brown (eds.). *The Routledge Companion to Thought Experiments*. London and New York: Routledge, 512–525.
- Egan, D. 2016. “Literature and Thought Experiments.” *The Journal of Aesthetics and Art Criticism* 74: 139–150.
- Elgin, C. Z. 2014. “Fiction as Thought Experiment.” *Perspectives on Science* 22: 221–241.
- Gendler, T.S. 2010. *Intuition, Imagination and Philosophical Methodology*. Oxford: Oxford University Press.
- Hume, D. 1999. *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press.

- Kind, A. 2013. "The Heterogeneity of the Imagination." *Erkenntnis* 78: 141–159.
- Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
- Meynell, L. 2014. "Imagination and Insight: A New Account of the Content of Thought Experiments." *Synthese* 191 (17): 4149–4168.
- Meynell, L. 2018. "Images and Imagination in Thought Experiments." In M. T. Stuart, Y. Fehige and J. R. Brown (eds.). *The Routledge Companion to Thought Experiments*. London and New York: Routledge, 498–511.
- Miščević, N. 1992. "Mental Models and Thought Experiments." *International Studies in the Philosophy of Science* 6: 215–226.
- Miščević, N. 2007. "Modelling Intuitions and Thought Experiments." *Croatian Journal of Philosophy* 7: 181–214.
- Molinari, D. 2020. "Spoiler Alert! Unveiling the Plot in Thought Experiments and Other Fictional Works." *Argumenta* 6: 81–97.
- Murphy, A. 2020a. "Towards a Pluralist Account of the Imagination in Science." *Philosophy of Science* 87: 957–967.
- Murphy, A. 2020b. "The Aesthetic and Literary Qualities of Scientific Thought Experiments." In M. Ivanova and S. French (eds.). *The Aesthetic of Science. Beauty, Imagination and Understanding*. London and New York: Routledge, 146–166.
- Murphy, K. M. 2004. "Imagination as Joint Activity: The Case of Architectural Interaction." *Mind, Culture, and Activity* 11: 267–278.
- Murphy, K. M. 2005. "Collaborative Imagining: The Interactive Use of Gestures, Talk, and Graphic Representation in Architectural Practice." *Semiotica* 156: 113–145.
- Nersessian, N. J. 1992. "In the Theoretician's Laboratory: Thought Experimenting as Mental Modeling." *Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2: 291–301.
- Nersessian, N. J. 2018. "Cognitive Science, Mental Modeling, and Thought Experiments." In M. T. Stuart, Y. Fehige and J. R. Brown (eds.). *The Routledge Companion to Thought Experiments*. London and New York: Routledge, 309–326.
- Norton, J. D. 2004. "Why Thought Experiments Do Not Transcend Empiricism." In C. Hitchcock (ed.). *Contemporary Debates in the Philosophy of Science*. Oxford: Blackwell, 44–66.
- Norton, J. D. 2013. "Chasing the Light: Einstein's Most Famous Thought Experiment." In M. Frappier, L. Meynell and J. R. Brown, (eds.). *Thought Experiments in Philosophy, Science and the Arts*. London and New York: Routledge, 123–140.
- Parfit, D. 1984. *Reasons and Persons*. New York: Oxford University Press.
- Salis, F. 2020. "Learning Through the Scientific Imagination." *Argumenta* 6: 65–80.
- Salis, F. and Frigg, R. 2020. "Capturing the Scientific Imagination." In A. Levy and P. Godfrey-Smith (eds.). *The Scientific Imagination*. Oxford: Oxford University Press, 17–50.
- Searle, J. R. 1980. "Minds, Brains and Programs." *The Behavioral and Brain Sciences* 3: 417–457.
- Sorensen, R. A. 2019. "Smartfounding: Four Grades of Resistance to Thought Experiments." *Topoi* 38: 791–800.

- Stuart, M. T. 2020. "The Productive Anarchy of Scientific Imagination." *Philosophy of Science* 87: 968–978.
- Szanto, T. 2017. "Collective Imagination. A Normative Account." In M. Summa, T. Fuchs and L. Vanzago (eds.). *Imagination and Social Perspectives. Approaches from Phenomenology and Psychopathology*. London and New York: Routledge, 223–244.
- Thomson, J. J. 1971. "A Defense of Abortion." *Philosophy and Public Affairs* 1: 47–66.
- Vidmar, I. 2013. "Thought Experiments, Hypotheses, and Cognitive Dimension of Literary Fiction" *Synthesis Philosophica* 28: 177–193.
- Walton, K. L. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge: Harvard University Press.
- Willée, A. 2019. "Thought Experiments as a Narrative Genre." In F. Bornmüller, J. Franzen and M. Lessau (eds.). *Literature as Thought Experiment? Perspectives from Philosophy and Literary Studies*. Paderborn: Wilhelm Fink, 83–96.
- Wittgenstein, L. 2009. *Philosophical Investigations*. Oxford: Blackwell.
- Wittgenstein, L. 2018. "Remarks on Frazer's *The Golden Bough*." In G. da Col and S. Palmié (eds.). *The Mythology in Our Language*. Chicago: Hau Books, 29–75.