Thought Experiments
Imagination in Practice

Sofia Pedrini

> The scientist must lack prejudice to a degree where he can look at the most "self-evident" facts or concepts without necessarily accepting them, and, conversely, allow his imagination to play with the most unlikely possibilities.
> *Selye, qtd in Kuhn 226*

Abstract *This paper proposes that thought experiments (TEs) should be understood as imaginative practices based on narrative construction. I consider them practices because they require active collaboration between the author and an engaged reader—essential to the TE's execution and success. These practices are also imaginative since the reader, guided by the author's narrative, constructs an imaginary scenario that brings the author's theoretical perspective into view. The purpose of this scenario is to prompt the reader to realize or intuit something crucial about the author's theoretical position. Rather than merely presenting a perspective, a TE supports a thesis or claim through this imaginatively shared theoretical perspective. Additionally, I describe the reader's participation as a form of immersion within the TE's imagined scenario. Finally, by drawing an analogy with real-world experiments, I argue that TEs function as a social imaginative practice, fostering discussion and dialectical exchange within specific scientific communities.*

Keywords *Thought Experiments; Imagination; Immersion; Scientific Practices; Social Practices; Epistemology; Phenomenology*

1. Introduction

The notion of thought experiment refers to a particular kind of experiment performed "in the mind." Philosophers and scientists have used thought experiments (hereafter TEs) in their writings to support their theories, strengthen their arguments, or make a case for a particular claim (see e.g. Williamson). Some of the most famous TEs in philosophy are Plato's *Gyges' Ring*, Descartes's *Evil Genius*, Locke's *The Prince and the Cobbler*, Nozick's *Experience Machine*, Putnam's *Twin Earth*, Jackson's *Knowledge Argument*, Thomson's *Ailing Violinist*, and Chalmers's *Zombies*. Some influential scientific TEs include Galilei's *Falling Bodies*, Stevinus's *Inclined Plane*, Einstein's *Moving Trains* (*and Elevators*), Maxwell's *Demon*, and Newton's *Bucket*.[1] Besides philosophy and the physical sciences, TEs have been employed in various fields, from politics and economics to the fine arts (e.g. Frappier, Meynell and Brown; Brown and Fehige). Despite this variety, TEs are easily recognizable as a representation of a counterfactual or fictional situation that the authors insert alongside, and in support of, their arguments.

In this paper, I propose that TEs ought to be understood as imaginative practices based on narrative construction. I first present the structure of TEs and argue that the imaginary scenario is constructed to prompt the reader to realize or intuit something crucial about the author's theoretical perspective. In my view, the imaginary scenario is necessary to bring the author's theoretical perspective into the thought experimenter's view, and yet, I argue, the TE does not merely show us a particular theoretical perspective, but aims at supporting a thesis or claim using that perspective.

Moreover, I analyze the dynamic between the author and the reader of a TE. As in our usual engagement with fiction, this dimension can be characterized as a specific norm-governed dynamic between the author and the reader: there are some "rules" that the author of a TE implicitly communicates through the narrative on how the reader must construct the imaginative scenario to properly engage in that TE. After constructing and immersing in the imaginary scenario, the reader must reason it out: to be performed, the TE requires the reader *to do something*. The idea is that, while reading the TE and reasoning on the imaginary scenario, the reader is performing or recreating the thought experiment—i.e., they are the thought experimenter.

1 See Gendler, Thought Experiment chap. 1.

Finally, I propose that TEs can be considered a broader social practice, as they promote discussion and dialectical exchange within a specific scientific community. Just as we wouldn't consider an experimental research finished or complete when we gain data, as the data must be interpreted, discussed, and where possible replicated, we should not consider a TE-based research finished or complete when we (as readers) gain an intuition, for we still must interpret the intuition, discuss our interpretations, and replicate the TE where possible.

I proceed in the following way. In §2, I examine the structure of TEs through Gendler's perspective and those of Binini, Huemer, and Molinari, using Jackson's *Knowledge Argument* as a case study. I choose these views as they best illustrate the connection between imaginary scenarios, argumentation, and reasoning. In §3, I argue that "imaginary" in TEs refers not to the impossibility of the scenarios but to the fact that they are created and reasoned through imagination, as demonstrated by Galilei's *Falling Bodies* TE. In §4, I analyze the different types of imagination and argue that imaginative immersion is crucial for TEs. This leads to an exploration of the author-reader dynamic in TEs in §5.[2] Finally, in §6, I argue that TEs are practices embedded within scientific communities and conclude that they are also *social* practices, thereby shedding light on the role of TEs—and, consequently, imagination—in scientific and philosophical debates.

2. The Structure of Thought Experiments

To illustrate how the imaginary scenario of a TE is necessary to bring the author's theoretical perspective into the thought experimenter's view, I analyze one of the most influential TEs in contemporary philosophy, Frank Jackson's *Knowledge Argument*, through the lens of Gendler's tripartite structure of TE narrative. According to Gendler, in a TE:

> (1) An imaginary scenario is described. (2) An argument is offered that attempts to establish the correct evaluation of the scenario. (3) This evalua-

2 The standard way of engaging with a TE is through reading. One might argue, however, that my points should not be limited to this way, as a TE could also be presented orally to an audience. For ease of exposition, I will focus on the author-reader dynamic in the following discussion, though I believe my account extends to an 'author-audience' dynamic as well.

> tion of the imagined scenario is then taken to reveal something about cases beyond the scenario (*Thought Experiment* 21).[3]

In his 1982 paper "Epiphenomenal Qualia," Jackson writes:

> Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like 'red', 'blue', and so on. […] What will happen when Mary is released from her black and white room or is given a color television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. (130)

Jackson is here building an argument against physicalism, that is, a metaphysical reductionist position according to which everything that exists is fundamentally constituted by physical entities. Jackson's argument aims to show that there are phenomenal facts, i.e., first-person experiential facts—such as the experience of seeing a color—and these facts are not reducible to physical facts, i.e., third-person, impersonal facts. To do so, the author describes the imaginary scenario of Mary who has spent her life in a black-and-white room, while knowing all the physical facts about colors; at one point, Mary leaves the room and sees colors for the first time, e.g., red.[4] This imaginary scenario conveys to the reader the following intuition: Mary learns something new when she first sees the color red. Even with an exhaustive knowledge of physical facts about colors, Mary in the black-and-white room would still lack knowledge of the phenomenal fact of "seeing red." The tripartite structure of this TE is the following: (1) Mary's scenario is described; (2) The evaluation of the scenario is offered: Mary learns something new when she sees colors for the first time; (3) This evaluation is taken to reveal that (a) phenomenal facts are not deducible

3 See also Wiltsche's analysis of the 3 stages of a TE.

4 For a proper analysis of the relevance of each single TE, it must be contextualized in the philosophical or scientific debate their author was engaging in; in Jackson's case, the Knowledge Argument against Physicalism. For the purposes of this paper, I gloss over more specific conclusions drawn from this TE and the theoretical and philosophical debates this argument is involved with, to rather focus on how it works. See also Chalmers 103–104.

from or implied by physical facts, or phenomenal facts are not physical facts, ergo (b) physicalism is false (not everything is reducible to physical facts).

Similarly, another interesting view proposes that TEs are "games" of theoretical perspective-taking (Binini et al. 1). In this view, TEs are "particularly apt to express a specific theoretical perspective through the use of imagination" (1), as they "invite the readers to engage in 'games of perspectives'" (7). Binini et al. draw on Walton's theory of fiction, particularly his central notion of make-believe games, which suggests that imagination is not merely a private experience but can also be conceived as "an activity that can be shared with others and involves a *normative dimension*" (4, emphasis mine)—as seen in children's pretend play or actors embodying a role. From this view, TEs can be understood as specific forms of make-believe games. Mary's TE is a make-believe game in which we, as readers, adopt Jackson's theoretical perspective on phenomenal facts. This imaginative practice leads us to take Mary's scenario as supporting Jackson's claim, ultimately serving as a counterexample to physicalism.

TE is not merely a question of showing the author's perspective through an imaginary scenario, but necessarily involves a "process of reasoning carried out [by the thought experimenter] within the context of a well-articulated imaginary scenario to answer a specific question about a non-imaginary situation" (Gendler, *Thought Experiment X*). A good TE thus is one that effectively presents the author's theoretical perspective, making it compelling to the reader using imagination and the intuition drawn from the imagined scenario. Through Mary's scenario, we can adopt Jackson's theoretical perspective. Yet, the TE does not merely show us a particular theoretical perspective but aims to support the author's thesis or claim.

3. The Scenario is *Imaginary*

An interesting and controversial aspect of TEs is the nature of the scenarios described by authors to support their theories or claims: they are imaginary. It is important to clarify that "imaginary" does not imply that these scenarios are always impossible.[5] Although TEs often concern cases that presumably cannot

5 The notions of possibility and impossibility can be categorized into three dimensions: nomologically (i.e., physically, biologically, chemically, etc.) possible and impossible, logically (or conceptually) possible and impossible and metaphysically possible and

occur in our world, they can also describe scenarios that could have occurred, i.e., counterfactual scenarios.

An example of a TE that addresses a potentially real-world case is Galileo Galilei's *Falling Bodies* TE, which aimed to disprove a specific prediction of Aristotelian physics regarding the laws of motion for falling bodies. This TE aims at making a case against Aristotelian physics according to which heavy bodies would fall faster than lighter ones, thereby supporting Galilei's view that bodies of different weights fall with the same acceleration in vacuo (see De Angelis xvii). In a central passage of his 1638 "Two New Sciences", through the character of Salviati, Galilei states that:

> [E]ven without further experiment, it is possible to prove clearly, by means of a short and conclusive argument, that a heavier body does not move more rapidly than a lighter one provided both bodies are of the same material and in short such as those mentioned by Aristotle. (De Angelis 44)[6]

Further on, to support his view, Galilei describes an imaginary scenario as follows: imagine standing atop the Tower of Pisa, dropping two stones of different sizes, and then imagining them joined together by a rope as they fall.

> If then we take two bodies whose natural speeds are different, it is clear that on uniting the two, the more rapid one will be partly retarded by the slower, and the slower will be fastened by the swifter […]. But if this is true, and if a large stone moves with a speed of, say, eight grades, while a smaller one moves with a speed of four, then when they are united, the system will move with a speed less than eight; but the two stones when tied together make a stone larger than that which before moved with a speed of eight. Hence the heavier body moves with less speed than the lighter; an effect which is contrary to your supposition. Thus, you see, it can't be true that the heavier mobile moves faster than the lightest. (De Angelis 44)

Through this TE, Galileo shows that, if we assume Aristotle's physics, we arrive at a contradiction, and the way out is to "infer therefore that large and small

---

impossible (see Gendler and Hawthorne). In a TE, the imaginary scenario does not need to be impossible in any of these respects.

6 The following quotations are taken from De Angelis' translation of Galileo Galilei's Two New Sciences for Modern Readers. For simplicity I will refer to the pages in this translation.

bodies move with the same speed provided they have the same specific weight" (45).[7] What interests us is that in this TE, the imaginary scenario could have been realized in a real experiment; Galileo "could have performed the experiment" (Brown, "Thought Experiments Since" 3). However, he chose to conduct it "in the mind," arguably for the sake of simplicity and practicality. The strength of the TE lies in the argumentation that arises from reasoning about the imaginary scenario under consideration. Here, "imaginary" refers to the creation of the scenario "in our mind," with the reasoning process conducted in imagination, which serves as "the medium in which TEs are performed" (Wiltsche 345). Thus, the role of imagination in a TE warrants further exploration.

4. Thought Experiments as *Imaginative* Practices

Imagination is a widely discussed concept in contemporary philosophical debates, yet it is challenging to define, with philosophers disagreeing on its phenomenology—specifically, what precisely happens when one imagines something (see Kind). Nevertheless, there is general agreement that imagination comes in different types or kinds (see Dokic and Arcangeli). What we usually refer to when considering imagination is what philosophers call sensory imagination, which involves mental imagery. When I imagine a red apple, I form a mental image of a red apple. Mental imagery can be multimodal, that is, it occurs in different sensory modalities beyond vision. I can imagine seeing an

7 Gendler reconstructs the TE as follows: "Imagine that a heavy and a light body are strapped together and dropped from a significant height. What would the Aristotelian expect to be the natural speed of their combination? On the one hand, the lighter body should slow down the heavier one while the heavier body speeds up the lighter one, so their combination should fall with a speed that lies between the natural speeds of its components. (That is, if the heavy body falls at a rate of 8, and the light body at a rate of 4, then their combination should fall at a rate between the two (cf. Galilei 1638/1989: 107).) On the other hand, since the weight of the two bodies combined is greater than the weight of the heavy body alone, their combination should fall with a natural speed greater than that of the heavy body. (That is, if the heavy body falls at a rate of 8 and the light body at a rate of 4, their combination should fall at a rate greater than 8.) But then the combined body is predicted to fall both more quickly, and more slowly, than the heavy body alone (cf. Galilei 1638/1989: 107–8). The way out of this paradox is to assume that the natural speed with which a body falls is independent of its weight: 'both great and small bodies … are moved with like speeds' (Galilei 1638/1989: 109)" (Intuition, Imagination 27–28).

apple, but I can also imagine (or form a mental image of) the sound of a piano sonata, the smell of a flower, the punch of a friend, or the taste of melon. These are all cases of *sensory* imagination (see Nanay). This type of imagination does not appear to be necessary for every kind of TE, or at least, it seems that it does not significantly affect the results of a TE if one vividly visualizes the Tower of Pisa and the two stones. It seems that the power of a TE derives from something beyond this.

There's another kind of imagination discussed in contemporary debates, *propositional* imagination, that need not involve mental imagery. Some examples are imagining that all wars on earth cease or imagining that my next paper will be extremely influential. This kind of imagination involves "merely imaginatively representing that something is the case, similar to merely entertaining or assuming a proposition" (Myers 3252). Yet, in TE we are doing more than merely entertaining some propositions. As we saw in the previous section, we are engaged and actively participate in a TE. Binini et al. draw on the notion of games of make-believe to describe this active engagement, noting that "players tend to actively participate in the game of make-believe" (Binini et al. 5) by employing first-person imaginative engagement. To clarify the nature of this imaginative experience, it is necessary to analyze this specific mode of engagement. According to Binini et al., this engagement can take two forms:

> [O]ne can [1] imagine oneself in the shoes of others and identify with (one of) the characters, or [2] imagine to observe the scenarios described from another (e.g., the narrator's) point of view. (Binini et al. 5)

The first form of first-person imaginative engagement [1] corresponds to *empathic* imagination which is the form of imagination we employ when we imagine being another person (see Smith). More precisely, it is imagining being the other person "from the inside" or from a first-person perspective. Interestingly, one may claim that some kind of empathic imagination is fundamental for performing a TE. Indeed, many TE narratives—mainly the philosophical ones—seem to be constructed to make the reader empathically imagine being, i.e., take the first-person perspective of the main character. In Mary's case, it seems like Jackson wants us to take Mary's perspective, to imagine being in her shoes, to understand *what it is like* to be in those circumstances described by the

imaginative scenario.[8] Is it necessary, though, to take the perspective of a character to perform a TE? The reply is negative: it is not necessary to empathically imagine being a fictional character of a TE in order to perform it, as it does not happen in Galileo's TE and many other TEs.

What, then, is the distinctive imaginative experience involved in all TEs? The answer lies in what Binini et al. identify as the second form of first-person imaginative engagement [2]: imagining oneself as observing the scenario from the narrator's perspective, thereby adopting the thought experimenter's stance. This form of engagement is best understood as *imaginative immersion*. According to Wiltsche, immersion "consists in the active bracketing and/or modification of certain parts of our background knowledge" (354), enabling us to pretend that the imaginary scenario is actually the case. In TEs, we immerse ourselves in the scenarios described by the author, going beyond merely entertaining a proposition. We engage with these scenarios *as if* they were real, reasoning about their appropriate descriptions, consequences, and effects. By following the trajectory outlined in the narrative, we explore the implications of the imaginary scenario as though it was real.[9]

Immersing oneself in a TE does not necessarily involve sensory imagination nor empathic imagination with a character of the TE and it seems to mean more than just propositionally imagining the scenario. Immersing oneself in the imaginary scenario means taking it as if *it were the case* (see Chasid). Moreover, through the narrative, the reader is "called" to take or adhere to—even only momentarily—the author's theoretical view on the scenario.

The presence of an imaginary scenario and our necessary immersion in it highlight how reductionist accounts, such as those of Norton and Williamson—who argue that the cognitive value of a TE lies solely in its argument, reducing TEs to mere arguments while dismissing the rest as mere ornamentation, especially regarding imagination as a non-essential embellishment—fail to capture the essential scope of a TE. The next section

8 There are some TEs, notably Nozick's Experience Machine and Thomson's Ailing Violinist, where the author wants us to reflect on what we would do or how we would act if we were in those circumstances, that is, those TEs where the reader's sense of agency and their decision-making, with all its implications and consequences, are relevant for the TE itself.

9 This form of immersion is comparable to engaging with a book, movie, or theatrical performance, where we imagine the world, facts, characters, stories, and consequences. For a discussion of immersion in fiction, see Sartre's account, which argues that such immersion does not necessarily require the creation of mental images.

aims to better understand this scope by analyzing the dynamics between the author and the reader in a TE, which mirror the engagement we have with narratives—another distinct form of imaginative immersion. If we consider imaginative immersion as a form of imaginative experience, it becomes clear that imagination is central to TEs, challenging reductionist views.

5. The Dynamics Between Author and Reader

My analysis of the dynamics between the author and the reader of a TE rests on two elements, which help clarify the structure of TEs and allow me to argue that TE is a *practice* where the thought experimenter achieves conclusions. The two elements of the TE reader-author dynamics are:

A. The strategic directions of the author of the TE for constructing the imaginary scenario;
B. The engagement or active participation of the reader.

Let's start with the strategic directions the author gives to construct the imaginary scenario (A). The ability of the TE's author lies in the construction of a particular narrative that can make their own theoretical perspective available to the reader by making a case for their theory. How does this happen? As in our engagement with narratives, when we engage in a TE, our imagination does not freely wander but follows the narrator's instruction on what to imagine. The narratives guide the reader in a particular way, as Wiltsche explains:

> [B]y following my instruction to imagine [e.g.] Bart Simpson, you immersed yourself in the quasi-world of *The Simpsons* and thus accepted certain limitations to your imagining. The concept "Bart Simpson" contains what has become known to you and your epistemic community about the kind of object in question. (Wiltsche 352)

When we follow instructions on what to imagine, we "choose to immerse ourselves in a quasi-world by staying within the boundaries that are prescribed by the concepts" (353) and are "encapsulated" (356) in the TE narrative. We could decide to imagine something which is incompatible with the concept "Bart Simpson," or we could "inadvertently fail to stay within the boundaries of what

the concept prescribes" (352). In such cases, we fail to immerse ourselves into the "quasi-world" of *The Simpsons*.

Binini et al. also stress the importance of the rules that "guide the readers' imagination" for properly constructing the imaginative scenario, "if one would refuse to imagine the given scenario, one would not play the game in an authorized way" (5).[10] Following the TE narrative, we implicitly agree on following the author's direction to "fairly" play the game of theoretical perspective-taking. These directions on what to imagine are not different from what happens when we normally engage in narratives: the author describes the characters, events, or scenes we must imagine while reading fiction, a history book, or any other book.[11] What is peculiar about the particular kind of narrative of the TE is the *aim* of the practice itself, that is "[…] to bolster or test our theories or concepts" (Wilkes 10). According to Wilkes, the "purpose in hand" is crucial for understanding what counts as relevant or not in our imaginative scenario (10). The success of a TE lies then in the construction of a particular narrative which can make *relevant* parts of the author's theoretical perspective available to the reader. In TE's case, the narrative drives us to *properly* establish the phenomenon (Brown, "Thought experiments since" 4) under analysis in imagination.[12]

We can now analyze the second element (B) of the author-reader dynamic: the interesting feature of TEs is a particular engagement or constructive participation of its reader (Gendler, *Thought Experiment* XIII). This element is adequately captured by the notion of "game" used by Binini et al. to refer to TEs:

> These games require active engagement from the reader, who has to accept the author's invitation—as well as the rules of the game—and cognitively immerse themselves into the fictional scenario, which allows them to critically assess it and to partake in the dialectical exchange of perspectives. (7–8)

10 As Wiltsche claims, TE narratives instruct us about what we are supposed to imagine, but they "also give us information about what we are not supposed to imagine (Cf. Davies 35). In part this is done through the target-thesis that automatically narrows our focus to certain aspects and leaves out others. But it is also done through the determination of the quasi-world in which the TE must be embedded" (Wiltsche 358).

11 More must be said about the relation between our usual engagement with narratives and TEs than I can do here.

12 Wilkes stresses the importance of properly establishing a phenomenon in imagination for the reliability of TEs.

The reader is called to *follow* the TE narratives, to accept their rules to properly construct the imaginary scenario and reason or reflect on it. As Gooding suggests, the reader's "personal participation is essential: it is what makes a thought experiment an *experiment* rather than another form of argumentation" (281).

To sum up, TEs can be considered a practice as (1) they essentially involve a particular dimension between author and reader (in the next section, we will see that this "minimal" dimension becomes a full social dimension of TEs), and (2) the reader is called in and asked to actively engage in the TE. We can now fully understand why I consider TEs *imaginative* practices: they essentially require immersion in an imaginary scenario, i.e. they do not involve merely entertaining an imaginary situation, but they require an effort to take on some beliefs and suspend others, with the goal of testing, challenging or (simply) confirming our theories, beliefs, and opinions. We've seen so far that TEs require collaboration between author and reader, and a particular engagement from the part of the reader; next we explore how TEs aim at the "intuitive grasp" of the author's theoretical claim built on the imaginative scenario. However, in the following paragraph, I argue that a TE does not end with intuition, but that this practice is embedded in larger social practices and open to public discussion of the results, precisely as non-imaginative experiments—those performed in the lab—are.

6. A *Social* Imaginative Practice: Interpretation, Discussion, and Variants

TEs can be considered imaginative practices where the reader comes to realize something about the relevant claim of the author's theoretical perspective, by immersing in and reasoning out the imaginary scenario. The results of a TE have been famously described by Daniel Dennett's definition of TEs as "intuition pumps" (*Elbow Room* 12, 17–18): TEs aim at pumping the relevant intuition in the reader. The analogy with actual experiments allows me to consider the *intuition* of a TE as analogous to the *data* of an actual experiment: actual experiments aim at gaining data, analogously TEs aim at gaining or "pumping" intuitions.[13]

13 The analogy with actual experiment is widely discussed in contemporary philosophical debates on the topic, and is already found in Mach.

> Intuition is neither a doxastic attitude, such as a belief or judgement, nor a mere tendency to form such an attitude, but rather a presentation: a conscious state or event that, like perceptual experience, directly and immediately presents the world as being a certain way (Bengson 708).[14]

When we have an intuition, it *seems* to us that things are the way we intuit them.[15] The intuition we gain in Mary's TE, for example, is that it seems to us that Mary learns something new when she goes out of the black-and-white room and sees the colors for the first time. This is the result we gain from Mary's TE.

We can understand how the main components of a TE, i.e., imagination and intuition, are related to each other in the following way: through the construction of the TE imaginative scenario, the author aims at making the relevant intuition compelling to the reader who must actively engage in the TE. The author can give proper instructions (through the narrative) to pump the relevant intuition into the reader. Intuition can thus be considered the 'data' of TEs, subject to discussion in order to work out its implications. This aspect of TEs is derivable from the analogy with non-imaginative scientific experiments (see Brown and Fehige) and is central to my understanding of TEs as a social practice: just as we would not consider experimental research concluded when we get the data, similarly, we should not consider a TE-based research concluded when we (as readers) get the relevant intuition. As Nersessian writes, "a thought experimental outcome, just as a real-world experimental outcome, needs to be interpreted and usually investigated further" (310).

Considering Mary's case again, Dennett is not convinced by Jackson's interpretation of the TE's result/intuition and proposes a "counter thought experiment" (see Brown, "Counter Thought Experiments") to disprove Jackson's TE. Dennett writes:

---

14 Intuition is a problematic notion as it is claimed to be culturally and socially determined. In this paper, I set aside the epistemological problem of intuition, i.e., whether intuitions can be reliable, and how they lead to knowledge, if they even can. Even though it's a central issue especially in meta-philosophy (the philosophical investigation on the methodology of philosophy itself), I'm interested here in understanding what intuition is, what kind of mental state it is—what happens when we have an intuition. See also Chudnoff for an analysis of intuition, and Hopp on the role of intuitions in a TE.

15 Like perception, intuition is fallible: our intuitions can be mistaken.

> And so, one day, Mary's captors decided it was time for her to see colors. As a trick, they prepared a bright blue banana to present as her first color experience ever. Mary took one look at it and said, "Hey! You tried to trick me! Bananas are yellow, but this one is blue!" Her captors were dumfounded. How did she do it? "Simple," she replied. "You have to remember that I know everything—absolutely everything—that could ever be known about the physical causes and effects of color vision. So of course before you brought the banana in, I had already written down, in exquisite detail, exactly what physical impression a yellow object or a blue object (or a green object, etc.) would make on my nervous system. So I already knew exactly what thoughts I would have (because, after all, the 'mere disposition' to think about this or that is not one of your famous qualia, is it?). I was not in the slightest surprised by my experience of blue (what surprised me was that you would try such a second-rate trick on me). I realize it is *hard for you to imagine* that I could know so much about my reactive dispositions that the way blue affected me came as no surprise. Of course it's hard for you to imagine. It's hard for anyone to imagine the consequences of someone knowing absolutely everything physical about anything! (*Consciousness Explained* 399–400)

Dennett's point is that if Mary knows all the physical facts, she would also know what a yellow banana looks like because, if she is omniscient about colors, she must be able to know a banana looks like (i.e., we cannot accept Jackson's conclusion, and should instead conclude that Jackson assumed phenomenal facts cannot be derived from physical facts). To prove his point, Dennett proposes this variant of the TE where the resulting intuition no longer concerns the relationship between phenomenal and physical facts. Instead, it shifts to Mary's response to her surprised captors: "It is *hard for you to imagine*" being omniscient about colors (and about anything). If we can't fully imagine what that would entail, we can't have a reliable intuition about what Mary feels when she steps out of her room—meaning Jackson's conclusion is not reliable.

This Jackson-Dennett exchange illustrates how TEs serve as arenas for discussion. As Brown and Fehige note, TEs can be "rethought" (9), opening space for proposing variants.[16] When we gain an intuition, we must interpret

16 "A thought experiment proposed by one scholar invites others to explore a certain fictional scenario to arrive at a clear-cut conclusion—which might nevertheless run against the interlocutor's or the reader's intuitions. When participating in the game, these interlocutors will likely propose variations, be it because they fill in the blanks with details that run against the author's intentions, or because they vary the scenario by revising some of the principles of generation. In this way, they can critically adapt

it, discuss those interpretations, and propose alternative ones. This reveals that, much like a non-imaginative experiment, the practice of TEs is embedded within the relevant scientific community. As such, TEs are a *social* practice with the specific function of fostering dialectical exchange within a scientific or philosophical community (cf. Binini et al. 20).

7. Conclusions

In this paper, I have argued that TEs should be understood as imaginative practices grounded in narrative construction. The narrative of a TE describes an imaginary scenario in which the reader must immerse themselves and reason to arrive at the relevant intuition that sheds light on the author's theoretical perspective. My analysis has highlighted the central role of imaginative immersion in the execution and success of TEs, thus emphasizing the role of imagination in epistemic practices—particularly in engaging with philosophical ideas and generating deeper insights into the theories of scientists and philosophers.

Furthermore, I have argued that TEs are social practices, as they foster discussion and dialectical exchange within scientific communities. They serve as arenas where intuitions can be discussed, revisited, and rejected. TEs are also critical spaces where we can test the powers and constraints of our imagination. To what extent can we rely on the intuitions we derive from imagination? What scenarios are imaginable, and which are too difficult to imagine? How much immersion is necessary to reach a reliable conclusion? How can we better facilitate discussions of the results and criticisms of the premises?

Many questions remain open, and much work is needed to fully understand the implications and applications of TEs across various fields, including economics and politics. For instance, one pressing question is whether imagination—and TEs—hold social power in promoting alternative models of society, economics, or relations between states.

---

the scenario to their own theoretical needs and dialectical purposes" (Binini et al. 6). See Molinari for an analysis of the role of the imagination in scientific practice, where "clashes between imaginers" play a crucial role in the dialectic exchange between scientists or philosophers. Following the analogy of the actual experiment, one could argue that variants of TEs correspond to replications of the experiment.

By revealing the intrinsic connection between imagination and intuition in TEs, my analysis aims to pave the way for further research into the relationship between theoretical understanding and the practical uses of imagination. Moreover, through the analogy with non-imaginative experiments, I argue that, although our imaginative scenarios and the intuitions they elicit are imperfect, we can strive for a better understanding through ongoing discussion. This makes TEs, like any scientific experiment, a social enterprise that aims to expand our collective understanding.

Acknowledgments

I would like to thank the German Research Foundation DFG which supported this project in the context of funding the Research Training Group "Situated Cognition" (GRK 274877981). Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project Number GRK 274877981.

Works Cited

Bengson, John. "The Intellectual Given." *Mind*, vol. 124, no. 495, 2015, pp. 707–760.

Binini, Irene, Wolfgang Huemer and Daniele Molinari. "A Game of Perspectives: On the Role of Imagination in Thought Experiments." *Erkenn*, 2024.

Brown, James Robert. "Thought experiments since the scientific revolution." *International Studies in the Philosophy of Science*, vol 1, no. 1, 1986, pp. 1–15.

—. "Counter Thought Experiments", *Royal Institute of Philosophy Supplement*, vol. 61, no. 82, 2007, pp. 155–177.

Brown, James R., and Yiftach J. H. Fehige. "Thought Experiments." *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta and Uri Nodelman, 2012, pp. 135–142.

Chalmers, David J. *The Conscious Mind: In Search of a Fundamental Theory*, New York and Oxford, Oxford UP, 1996.

Chasid, Alon. "Imaginative immersion, regulation, and doxastic mediation." *Synthese* vol. 199, 2021, pp. 7083–7106.

Chudnoff,Elijah. *Intuition*. Oxford, Oxford UP, 2013.

De Angelis, Alessandro. *Galileo Galilei's "Two New Sciences" for Modern Readers*. Cham, Springer International Publishing, 2021.

Dennett, Daniel. C. *Elbow Room. The Varieties of Free Will Worth Wanting*. Oxford, Clarendon Press, 1984.

—. *Consciousness Explained*. New York, Little Brown, 1991.

Descartes, René. *Meditations on first philosophy. With selections from the objections and replies*. Translated by Michael Moriarty, Oxford, Oxford UP, 2008, p. 1641.

Dokic, Jérôme, and Margherita Arcangeli. "The Heterogeneity of Experiential Imagination." *Open MIND*, edited by Thomas K. Metzinger and Jennifer M. Windt, Frankfurt am Main, MIND Group, 2015.

Frappier, Melanie, Letitia Meynell and James Robert Brown, editors. *Thought Experiments in Science, Philosophy, and the Arts* (1st ed.). London, Routledge, 2012.

Gendler, Tamar. S. *Thought Experiment: On the Powers and Limits of Imaginary Cases*. Garland Press, New York, 2000.

—. *Intuition, Imagination, and Philosophical Methodology*. Oxford UP, New York, 2010.

Gendler, Tamar. S. and John Hawthorne, editors. *Conceivability and Possibility*. New York, Oxford UP, 2002.

Gooding, David. C. "What is experimental about thought experiments?" *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 2, Chicago, U of Chicago P, 1992, pp 280–290.

Hopp, Walter. "Experiments in Thought." *Perspectives on Science*, vol. 22, no. 2, 2014, pp. 242–263.

Horowitz, Tamara. and Gerald J. Massey, editors. *Thought Experiments in Science and Philosophy*. Lanham, Rowman and Littlefield, 1991.

Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly*, vol. 32, no. 127, 1982, pp. 27–36.

Kind, Amy. "The heterogeneity of imagination." *Erkenntnis*, vol. 78, 2013, pp. 141–159.

Kuhn, Thomas S. *The Essential Tension. Selected Studies in Scientific Tradition and Change*. Chicago, U of Chicago P, 1977.

Locke, John. *An essay concerning human understanding. 1690*. Oxford, Oxford UP, 1979.

Mach, Ernst. "Über Gedankenexperimente." *Zeitschrift für den physikalischen und chemischen Unterricht*, vol. 10, 1897, pp. 1–5.

Myers, Joshua. "The Epistemic Status of the Imagination." *Philosophical Studies*, vol. 178, no. 10, 2021, pp. 3251–3270.

Molinari, Daniele. "Thought experiments as social practice and the clash of imaginers." *Croatian Journal of Philosophy*, vol. 22, no.2, 2022, pp. 229–247.

Nagel, Thomas. "What Is It Like to Be a Bat?" *The Philosophical Review*, vol. 83, no. 4, 1974, pp. 435–50.

Nanay, Bence. *Mental Imagery Philosophy, Psychology, Neuroscience*. Oxford, Oxford UP, 2023.

Norton, John D. "Thought Experiments in Einstein's Work." *Thought Experiments in Science and Philosophy*, edited by Tamara Horowitz and Gerald Massey, Lanham, Rowman & Littlefield, 1991, pp. 129–148.

—. "Are Thought Experiments Just What You Thought?" *Canadian Journal of Philosophy*, vol 26, 1996, pp. 333–366.

Nersessian, N. "Cognitive Science, Mental Models, and Thought Experiments." *The Routledge Companion to Thought Experiments*, edited by Michael T. Stuart et al., London and New York, Routledge, 2018, pp. 309–326.

Nida-Rümelin, Martine, and Donnchadh O Conaill, "Qualia: The Knowledge Argument." *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, and Uri Nodelman, Stanford, Stanford University, Spring 2024.

Sartre, Jean-Paul. *Che cos'è la letteratura? Translated by L. Cogliati Arano, new augmented edition*. Milano, Il Saggiatore, 1966.

Smith, Steven R. "Empathic imagination and its limits". *Equality and diversity*, Bristol, Policy Press, 2011.

Sorensen, Roy A. *Thought Experiments*. Oxford, Oxford UP, 1992.

Walton,Kendall L. *Mimesis as Make-Believe: On the foundations of the representational arts*. Harvard UP, 1990.

Wilkes, Kathleen V. *Real People: Personal Identity without Thought Experiments*. Oxford, Oxford UP, 1988.

Williams, Bernard. *Problems of the self: philosophical papers 1956–1972*. Cambridge, Cambridge UP, 1976.

Williamson, Timothy. *The Philosophy of Philosophy*. 2nd ed., Hoboken, Wiley Blackwell, 2022.

Wiltsche, Harald A. "Phenomenology and Thought Experiments: Thought Experiments as Anticipation Pumps." *The Routledge Companion to Thought Experiments*, edited by Michael T. Stuart, et al., London and New York, Routledge, 2018, pp. 342–365.