

**Thought Experiments as Objects of Comparison<sup>1</sup>**

**The University of Tokyo, Kumpei YAMAMURO**

**1. Introduction**

As Timothy Williamson points out (Williamson 2022, 181), the method of thought experiments is the most conspicuous in analytic philosophy. However, varied views exist on thought experiments, and no consensus has been reached regarding what thought experiments are and how they serve philosophical inquiries. For example, James Robert Brown (1986, 1992, 2004, 2011) claims that some thought experiments are Platonic<sup>2</sup>, and John Norton (1991, 1996, 2004) contends that thought experiments can always be reconstructed as arguments; therefore, they are arguments in disguise.

Moreover, with the emergence of experimental philosophy, the importance of clarifying the nature and function of thought experiments cannot be overemphasized. Experimental philosophers have conducted empirical surveys to show that people's intuitions are significantly affected by apparently irrelevant factors, such as the subjects' cultural background, the order in which imaginary cases are presented, the environment in which the subjects' intuitions are elicited, and so on (see, e.g., Stich and Tobia 2017). Thus, experimental philosophy, albeit controversially, has provided much evidence to doubt the reliability of philosophers' intuitions. The plausibility and relevance of the challenges posed by experimental philosophy, however, depend on the answer to the question: What is a thought experiment? Suppose thought experiments are arguments in disguise, as Norton argues. In that case, the fact that people's intuitions are biased in many ways does not pose any hurdle to the plausibility of the method of thought experiments, just as the existence of the belief bias is irrelevant to the validity of an argument. However, if thought experiments are like real experiments, with the only difference being that they are

conducted in the brain, then the challenges posed by experimental philosophy can be critical because intuitions are then understood in an analogous way to data in real experiments.

Against the backdrop of such metaphilosophical concerns, I conceive of thought experiments as objects of comparison by arguing that Ludwig Wittgenstein's language games are a revised version of thought experiments. Wittgenstein's later philosophy contains many ingenious fictitious cases. However, the functions of those imaginary scenarios are not well understood, leading people to assume that Wittgenstein's language games are thought experiments despite his explicit denial. Nenad Mišćević, for example, asserts that Wittgenstein's imaginary scenarios count as thought experiments because they satisfy a standard definition of thought experiments. I begin by eliminating such a misunderstanding, delineating Wittgenstein's thoughts on language games and thought experiments.

This paper is organized as follows. Section 2 examines Mišćević's claim that Wittgenstein's imaginary scenarios are thought experiments. Section 3 presents Wittgenstein's thoughts on the concept of a thought experiment. Finally, Section 4 argues that Wittgenstein's method of language games is a revised version of thought experiments. Before concluding the paper, I suggest the advantage of the conception of thought experiments as objects of comparison over the prevalent one.

## 2. Mišćević on Wittgenstein's Scenarios

At the beginning of his paper, Mišćević quotes Wittgenstein on thought experiments:

What Mach calls a thought experiment is of course not an experiment at all. At bottom it is a grammatical investigation. (PR §1)

Then, Mišćević presents a standard definition of thought

experiments and asserts that Wittgenstein's scenarios are indeed thought experiments because they satisfy the definition, despite Wittgenstein's denial. Mišćević says:

Consider the definition of TE [= thought experiment] in Stanford Encyclopedia (due to J.R. Brown and Y. Fehige): "Thought experiments are devices of the imagination used to investigate the nature of things. Thought experimenting often takes place when the method of variation is employed in entertaining imaginative suppositions." The Builders scenario fits well; it is a device of the imagination used to investigate the nature of linguistic practices and to introduce the central idea of a language game. (Mišćević 2017)

What he calls the "Builders scenario" here is a language game Wittgenstein discusses in the beginning of *Philosophical Investigations*. Mišćević claims that the Builders scenario counts as a thought experiment because it satisfies the definition.

There are at least two problems with this claim. First, Wittgenstein's scenarios do not fit the definition. Second, Mišćević fails to grasp the point of Wittgenstein's assertion that thought experiments are not experiments; Wittgenstein is not saying that his scenarios should not be called "thought experiments" but that it is wrong to assume that thought experiments are pretty much like real experiments, with the only difference being that the former are conducted in the brain or the laboratory of the mind. In the remainder of this section, I deal with the first point; that is, I argue how Wittgenstein's imaginary scenarios do not satisfy the definition that Mišćević quotes. To this end, let us have a look at Wittgenstein's thoughts on philosophy and language games.

Wittgenstein states that "[p]hilosophy just puts everything before us, and neither explains nor deduces anything" (PI §126); we can, therefore, conclude that Wittgenstein would not employ language games to "investigate the nature of things."

Nevertheless, in the quote above, Mišćević equates “the nature of linguistic practices” with “the nature of things” so that the definition of thought experiments will resemble his understanding of *Philosophical Investigations*. Even then, Wittgenstein’s scenarios fail to fit the definition because he does not attempt to investigate “the nature of linguistic practices” in the first place. Wittgenstein maintains:

We want to establish an order in our knowledge of the use of language: an order for a particular purpose, one out of many possible orders, not *the* order. For this purpose we shall again and again *emphasize* distinctions which our ordinary forms of language easily make us overlook. (PI §132)

Wittgenstein does not try to establish *the* order to clarify the nature of linguistic practices. Instead, his language games are meant to illuminate our language *for a specific purpose*. In this sense, the orders he tries to establish are in no way absolute; they are relative to specific purposes. So what is the purpose for which an order is established? Wittgenstein says:

All explanation must disappear, and description alone must take its place. And this description gets its light — that is to say, its purpose — from the philosophical problems. (PI §109)

The purpose is to solve (or dissolve) philosophical problems. It is only insofar as it serves to solve philosophical problems that he tries to establish an order by means of language games. It is mistaken to assume that he tries to discover the nature of linguistic practices. Thus, Wittgenstein’s scenarios do not fit the definition.

One might argue that Wittgenstein’s fictitious cases can be regarded as thought experiments if we adopt a less restrictive

definition of the term. However, this line of response would trivialize Wittgenstein's criticism. We need to understand the rationale for his rejection.

### 3. Wittgenstein on Thought Experiments

Previously, I argued that Wittgenstein's scenarios do not satisfy the definition of a thought experiment that Mišćević refers to. Mišćević seems to assume that terming Wittgenstein's scenarios "thought experiments" or otherwise is a simple verbal matter. That may be why he gives few arguments to endorse the claim that Wittgenstein's scenarios are thought experiments. However, this is not true, as I claim below.

First of all, let us look at what Wittgenstein says about thought experiments.

What Mach calls a thought experiment is of course not an experiment at all. At bottom it is a grammatical investigation. (PR §1)

(Seemingly, I am performing 'thought-experiments'. Well, they're simply not experiments. Calculations would be much closer.) (LW I §519)

What is shown in these quotes is that Wittgenstein repudiates the idea of thought experiments because he thinks they are not experiments, and what we call thought experiments are much closer to "calculations" than to experiments. In other words, Wittgenstein is not so much opposed to what we call thought experiments being called "thought experiments" as he is opposed to them being considered experiments. Therefore, it is misguided to bring up a definition of thought experiments and argue that Wittgenstein's scenarios can be called "thought experiments" simply because they conform to it.

But what does he mean when he says that thought

experiments are not experiments but much closer to calculations? To answer this question, we need to examine Wittgenstein's discussions on the distinction between experiments and calculations.

According to Wittgenstein, normativity is key to differentiating calculations and experiments. Suppose that we make a calculation, say,  $25 \times 25$ , and get 625 one day and 624 the next day. We would naturally suspect that we made a mistake, at least in one of these calculations, which shows that we think that there must be right and wrong answers when it comes to calculations. In other words, normativity is intrinsic to calculations. At the same time, when we conduct experiments, we are concerned with what results we will obtain, not what the right results are<sup>3</sup>. So if the conditions of an experiment are fulfilled, "then we must accept whatever comes, as the result" (RFM 336). Obviously, this is not how we use calculations. Thus, treating calculations as experiments leads to a bizarre consequence:

[I]f a calculation is an experiment then the proposition that it yields such and such a result is after all the proposition that under such conditions this kind of sign makes its appearance. And if under these conditions one result appears at one time and another at another, we have no right to say "there's something wrong here" or "both calculations cannot be all right", but we should have to say: this calculation does not always yield the same result (*why* need not be known). (ibid.)

In most cases<sup>4</sup>, when making calculations we do not want to know how the subject will behave or what number a specific individual will arrive at in such and such conditions, which is where Wittgenstein thinks the distinction between experiments and calculations lies. If a calculation is an experiment, and the conditions are fulfilled, we have no right to say that the result is right or wrong. As Wittgenstein says: "The conditions of the

experiment don't include the result" (LFM 97).

However, when it comes to thought experiments, people freely say certain results or intuitive judgments are wrong or unreliable. For example, some proponents of experimental philosophy criticize the use of thought experiments and intuitions therein, and some philosophers who defend the method claim that laypeople's intuitions are unreliable because they lack expertise (see, e.g., Williamson 2022). Meanwhile, experimental philosophers have offered empirical evidence that even professional philosophers are not immune to various cognitive biases (see, e.g., Schwitzgebel and Cushman 2012) and argue that philosophers do not possess the kind of expertise they claim. In summary, people seem to be arguing about the right and wrong results of thought experiments. If thought experiments are experiments, then how can they be made sense of?

One way to make sense of them is to assume that philosophers are discussing the conditions of experiments when they seem to be arguing about the right and wrong results. Those who say some results are wrong and others are right must be taken as saying that certain conditions are not fulfilled.

But what are the conditions? They may be, for instance, such that the intuitions must be elicited from professional philosophers, or every detail of the thought experiment must be fully understood. If thought experiments include such conditions, people may be able to contend that some results are unreliable because the conditions are not fulfilled. For instance, those who say that experimental philosophers' challenges are irrelevant because they carried out surveys on the wrong subjects can be understood as complaining about the inadequacy with which experimental philosophers conducted the experiment.

Here, the problems are that (1) philosophers do not admit the divergence of intuitions even if the conditions are fulfilled, and (2) the meaning of the condition that every detail of the thought experiment must be fully understood is not clear. I discuss each of these below.

First, I have suggested above that the conditions of thought experiments may include one that the subjects should be philosophers. However, a homogeneous group of philosophers can also report divergence in intuitive judgments (e.g., Stich and Tobia 2017). In other words, even if the condition that the subjects should be philosophers is fulfilled, they would not cease talking about the right and wrong results, which shows that thought experiments are not treated as experiments. Philosophers think thought experiments allow for right and wrong results.

Second, some people might stipulate that subjects of thought experiments should understand the scenario *well* so that the results converge. In this case, however, normativity is built into the conditions of the experiment. As normativity is what differentiates experiments and calculations, this response must fail. Since the conditions of the experiment don't include the result, the criterion for understanding well should not be included in the conditions.

Therefore, it is unlikely that philosophers treat thought experiments as experiments. Philosophers discuss right and wrong results, but as we have seen above, it is difficult to interpret them as addressing the conditions of the experiment. Even if we call them "thought experiments," we do not treat them as such; we assume that the results are included in the conditions. What we call "thought experiments" are not experiments.

Then, how are we to understand Wittgenstein's imaginary cases? In the next section, I will argue that Wittgenstein's language games are a revised version of thought experiments, that is, thought experiments used not as experiments but as objects of comparison.

#### **4. Thought Experiments as Objects of Comparison**

Earlier, I illustrated Wittgenstein's claim that thought experiments are not experiments but closer to calculations. This

section deals with the positive characterization of what we call “thought experiments.” I first present Wittgenstein’s conception of language games as objects of comparison. Then, I argue that thought experiments should be taken as objects of comparison. Finally, I outline a possible significance of the revised conception of thought experiments.

Imaginary scenarios are ubiquitous in Wittgenstein’s later writings, and some people call them “thought experiments” just because they are fictitious. However, as we saw in the previous section, he is against the idea of thought *experiments*. What are the fictitious language games in his writings, if not thought experiments? How are they supposed to function? Wittgenstein explains the method of language games as follows:

Our clear and simple language-games are not preliminary studies for a future regimentation of language—as it were, first approximations, ignoring friction and air resistance. Rather, the language-games stand there as *objects of comparison* which, through similarities and dissimilarities, are meant to throw light on features of our language. (PI §130)

Although Wittgenstein criticizes the conception of imaginary scenarios as thought experiments, he recognizes the importance of constructing fiction in philosophy (cf. CV 74e). In this quote, he offers a positive characterization of language games as objects of comparison<sup>5</sup>. Wittgenstein’s language games are a revised version of thought experiments. That is, not only does he reject the conception of imaginary scenarios as thought experiments, but he proposes a viable way of using fictitious cases in philosophy. The fact that he refers to what is called a thought experiment as a grammatical investigation seems to endorse this claim (cf. PR §1), for “grammatical investigation” is another name for Wittgenstein’s later philosophy.

How does the conception of thought experiments as objects

of comparison connect with the claim that thought experiments are much closer to calculations? To answer this question, let us look at Wittgenstein's remarks about calculations.

If it is a calculation, we *adopt* it as a calculation—that is, we make a *rule* of it. We make the description of it the description of a *norm*—we say, “This is what we are going to compare things with.” (LFM 98)

Wittgenstein straightforwardly connects calculations with objects of comparison or standards of comparison.

Additionally, a mundane example can easily be offered because we always use calculations as objects of comparison in everyday life<sup>6</sup>. If you put two things in a bag and then add three but find only four items in the bag, you would suspect that you miscounted them or failed to put a thing into the bag. Here, the calculation “ $2 + 3 = 5$ ” serves as an object of comparison or a paradigm. You compare it with reality and assume you made a mistake. Note, however, that calculations do not apply to all experiences. If you add two drops to three drops of water, you will get one instead of five. So, it is crucial to remember that they are, after all, objects of comparison. Calculations or objects of comparison are not, as it were, responsible for reality; they are not preconceptions “to which reality *must* correspond” (PI §131).

Therefore, it is consistent to think Wittgenstein's language games are a revised version of thought experiments with his claim that thought experiments are more like calculations than experiments. As I argued above, calculations can also serve as objects of comparison. Additionally, what Wittgenstein says about the method of language games applies to calculations. He says a language game is not a preconception to which reality must correspond. Similarly, we can say that a calculation is not a preconception to which reality must correspond (recall the example of drops of water). Both language games and calculations are adopted because they are useful for some purposes.

What is wrong with the conception of thought experiments as experiments? Wittgenstein says that “we can avoid unfairness or vacuity in our assertions only by presenting the model as what it is, as an object of comparison” (PI §131), but he doesn’t provide further clarifications. So, I offer a possible consequence of taking fictitious scenarios as experiments, although the paper does not allow for a detailed discussion.

One obvious drawback is shown in the current metaphilosophical debates on the reliability of philosophical intuitions. My contention is that it is the conception of thought experiments as experiments that gives rise to the challenges of negative experimental philosophy. An easy way to see the point is to consider the fact that we do not cast doubt on the reliability of intuitions when we calculate, say,  $2 + 3 = 5$ <sup>7</sup>. It is irrelevant to the validity of calculations that we are prone to various biases or that calculations don’t hold for some sorts of objects. Likewise, we don’t take someone’s answering “ $2 + 3 = 6$ ” as a counterexample for a mathematical theory. If we make it perspicuous that thought experiments are not experiments, such doubts should evaporate.

One might argue that thought experiments are very much unlike calculations in that we have a variety of “intuitions” about an imaginary case while we have *the* answer to a given calculation; hence, it must be wrong to understand thought experiments as something close to calculations. My claim is that it is because we mistakenly take thought experiments as experiments that it seems we cannot discern “reliable intuitions.” In philosophical thought experiments, people’s intuitions count as evidence for or against philosophical claims. We think our intuitions need to be reliable enough to count as evidence, which in turn leads to experimental philosophical endeavors<sup>8</sup>. Thus, at the root of current philosophical debates about the reliability of intuitions lies the conception of thought experiments as experiments.

Note that I am not saying that experimental philosophy is entirely futile. Studies on the Knobe effect demonstrate the

significance of experimental philosophy (see, e.g., Knobe 2003; 2010). My claim is that by construing so-called thought experiments as objects of comparison, one can avoid the challenge posed by experimental philosophy as it stems from the conception of thought experiments as experiments<sup>9</sup>. However, this is just an outline. To delve deeper, we have to look at Wittgenstein's philosophy of mathematics, especially his view on the status of mathematical propositions. Due to the lack of space, the task should be tackled in future work.

## 5. Concluding Remarks

Wittgenstein denies that thought experiments are experiments, accentuating the distinction between experiments and calculations. According to him, what we call thought experiments are not experiments at all, but closer to calculations. Therefore, Wittgenstein's language games should not be considered thought experiments. They are objects of comparison that "throw light on features of our language" (PI §130).

I suggested a connection between experimental philosophy and the conception of thought experiments as experiments. However, due to the lack of space, I could only touch upon a possible connection. Future work is needed to show that the alternative conception of thought experiments as objects of comparison avoids the challenge of experimental philosophy, and that the conception continues to serve philosophical inquiries.

## Notes

1. This paper is sponsored by Japan Society for the Promotion of Science (22J20062).

2. “A *Platonic thought experiment* is a single thought experiment which destroys an old or existing theory and simultaneously generates a new one; it is *a priori* in that it is not based on new empirical evidence, nor is it merely logically derived from old data; and it is an advance in that the resulting theory is better than the predecessor theory.” (Brown 2011: 99)

3. There might be cases where we are inclined to say that certain results of an experiment are wrong (e.g., results that seem to indicate that neutrinos travel faster than light). However, even in such cases, it is not the results that we say are wrong but the conditions of the experiment.

4. Wittgenstein says that the act of calculating can function as an experiment in some cases. For example, if the purpose is to see whether a student has mastered multiplication or whether chalk will stand the strain, the act of calculating can count as an experiment (cf. RFM: 380; LFM: 93).

5. For the elucidation of how language games such as the Builders scenario are used as objects of comparison, see Kuusela 2019: Section 5.2.

6. It is obvious that calculations are not, in most cases, “meant to throw light on features of our language” (PI §130), but it does not follow that calculations are not objects of comparison. It is just that they are not the objects of comparison with our language.

7. I do not mean that we cannot cast doubt on the validity. Of course, this point is related to the so-called rule-following considerations, but I cannot deal with the connection here.

8. “If intuitions generated in response to thought-

experiments are supposed to be able to be used as reasons to accept or reject some philosophical claim, then we should be interested in studying the nature of the relevant intuitions. Experimental philosophy takes up this challenge, applying the methods of experimental psychology to the study of the nature of intuitions generated in response to thought-experiments.” (Alexander and Weinberg 2007).

9. If thought experiments could be interpreted as arguments as Norton tries to do, the problem of the reliability of intuition might be avoided. However, several problems exist with this claim. See, for example, Häggqvist 2009.

(3,795 words)

### References

- Alexander, J. and Weinberg, J. (2007). Analytic Epistemology and Experimental Philosophy. *Philosophy Compass*, 2(1), 56–80.
- Brown, J. R. (1986). Thought Experiments since the Scientific Revolution. *International Studies in the Philosophy of Science*, 1(1), 1–15.
- . (1992). Why Empiricism Won’t Work. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2, 271–279.
- . (2004). Why Thought Experiments Transcend Empiricism. in *Contemporary Debates in the Philosophy of Science*. edited by Christopher Hitchcock, Blackwell, 23–43.
- . (2011). *The Laboratory of the Mind: Thought Experiments in the Natural Sciences*. 2<sup>nd</sup> edition. Routledge.

Häggqvist, S. (2009). A Model for Thought Experiments. *Canadian Journal of Philosophy*, 39(1), 55–76.

Knobe, J. (2003). Intentional Action and Side Effects in Ordinary Language. *Analysis*, 63(3), 190–194.

———. (2010). Person as Scientist, Person as Moralist. *Behavioral and Brain Sciences*, 33(4), 315–329.

Kuusela, O. (2019). *Wittgenstein on Logic as the Method of Philosophy: Re-Examining the Roots and Development of Analytic Philosophy*. Oxford University Press.

Miščević, N. (2017). Wittgenstein's Thought-Experiments. *Conceptus*, 42(101–102), 20–40.

Norton, J. (1991). Thought Experiments in Einstein's Work. in *Thought Experiments in Science and Philosophy*. edited by Tamara Horowitz and Gerald Massey, Rowman & Littlefield, 129–148.

———. (1996). Are Thought Experiments Just What You Thought? *Canadian Journal of Philosophy*, 26(3), 333–366.

———. (2004). On Thought Experiments: Is There More to the Argument? *Philosophy of Science*, 71(5), 1139–1151.

Stich, S. and Tobia, K. (2017). Intuition and its Critics. in *The Routledge Companion to Thought Experiments*. edited by Michael T. Stuart, Yiftach Fehige, and James Robert Brown, Routledge, 369–384.

Schwitzgebel, E. and Cushman, F. (2012). Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-Philosophers. *Mind and Language*, 27(2), 135–153.

Williamson, T. (2022). *The Philosophy of Philosophy*. 2<sup>nd</sup> edition. Wiley-Blackwell.

Wittgenstein, L. (1975). *Philosophical Remarks*. edited from his posthumous writings by Rush Rhees, translated by Raymond Hargreaves and Roger White. Basil Blackwell. (PR)

———. (1978). *Remarks on the Foundations of*

*Mathematics*. 3<sup>rd</sup> edition, edited by G. H. von Wright, R. Rhees, and G. E. M. Anscombe, translated by G. E. M. Anscombe, Basil Blackwell. (RFM)

———. (1980). *Culture and Value*. edited by G. H. von Wright in collaboration with Heikki Nyman, translated by Peter Winch, The University of Chicago Press. (CV)

———. (1982). *Last Writings on the Philosophy of Psychology, volume 1*. edited by G. H. von Wright and Heikki Nyman, translated by C. G. Luckhardt and Maximilian A. E. Aue, Basil Blackwell. (LW I)

———. (1989). *Wittgenstein's Lectures on the Foundations of Mathematics Cambridge, 1939 from the Notes of R. G. Bosanquet, Norman Malcolm, Rush Rhees, and Yorick Smythies*. edited by Cora Diamond, The University of Chicago Press. (LFM)

———. (2009). *Philosophical Investigations*. translated by G. E. M. Anscombe, P. M. S. Hacker and Joachim Schulte, revised 4th edition by P. M. S. Hacker and Joachim Schulte, Wiley-Blackwell. (PI)