

THE END OF THE CASE? A METAPHILOSOPHICAL CRITIQUE OF THOUGHT EXPERIMENTS

Santiago A. VRECH

ABSTRACT: In this paper I carry out two tasks. First, I account for one of the distinctive uses of thought experiments in philosophy, namely, the fact that just *a* thought experiment is sufficient to confute a well-established theory. Secondly, I present three arguments to defend the claim that, at least in philosophy, we should remove thought experiments from our metaphilosophical toolkit. The central premise that motivates these arguments is the following: the very methodology of thought experiments permits to construct different scenarios in which philosophical theories are refuted *ad infinitum*.

KEYWORDS: thought experiments, imagination, experimental philosophy, metaphilosophy, cases, intuition

§ 1. Introduction

Philosophers use cases – or thought experiments – as a manner of refuting or supporting theories and analyses.¹ The methodological import of cases is such that some of the most renowned arguments in different branches of analytic philosophy consist precisely of thought experiments. For instance, due to Gettier cases we have reasons to believe that the analysis of ‘knowledge’ as a justified true belief does not hold; likewise, we might believe that descriptivism is incorrect because we judge that the thought experiments presented by Kripke sufficiently show the defects of the theory.

In the current metaphilosophical debate about thought experiments, it is usually maintained that an intuition needs to be elicited in order for thought experiments to have argumentative power. Indeed, “philosophers use intuitive judgements about cases as evidence for (or against) philosophical theories” (Andow 2020, 1). Thus, according to this picture, in considering a case² we come to non-

¹ Although the use of thought experiments is not limited to the purposes of refutation or validation. Davidson’s Swampan is an example of a thought experiment used for illustrative purposes, whereas trolley cases are instances of what Cohnitz calls “puzzle cases:” cases used to make us ponder (Cohnitz 2020, 102-103).

² In the metaphilosophical literature, “thought experiments” are also referred to as “the method of

inferentially perceive that the case is (or is not) an instance where the theory holds (or does not).³ In the light of this, the elicited intuitions are taken to be groundless – non-inferential – and *a priori*. It is because of these attributes that what Sydney Shoemaker has characterised as a cartesian conception of introspection can be applied to this context as well, since the use of intuitions

involves a strong form of the doctrine that mental entities are “self-intimating,” and usually goes with a strong form of the view that judgments about one’s own mental states are incorrigible or infallible, expressing a super-certain kind of knowledge which is suited for being the epistemological foundation for the rest of what we know (1994, 271).

Now, do we really need to rely on intuitions to explain the argumentative use of cases? And, perhaps more importantly, are we correctly describing our own argumentative use of thought experiments in this way? Many problems can be avoided with a negative answer to both questions. Indeed, the tortuous debate about the nature and use of intuitions can be avoided by conceiving of thought experiments as *reasons*. Why do we think that an internalist account of intentionality is incomplete? Because of a thought experiment – a twin earth scenario – that we wield as a reason. Why do we believe that the environment has any role to play in determining the aboutness of our concepts? Because of thought experiments! Why do we believe that knowledge is not a justified true belief? Because of cases! We use cases as reasons directly; there is no *philosophical* need to talk about intuitions to see the argumentative import that thought experiments have.

Reasons are important, and peculiarly so in philosophy. We put forward reasons and arguments for accepting or rejecting theories and analyses. Thought experiments are especially remarkable for the purpose of rejection; the philosophical damage that they cause is quite characteristic of them. So much so that, for example, when Gettier and Kripke published their cases, the theories against which they were directed came under attack immediately and were virtually overthrown. From this two points can be deduced: first, that just *a* case is taken to be enough to refute a well-established case, and secondly, that an imagined scenario has a striking refutative power. How is it possible for a case to dethrone an entire theory? Wherein lies this epistemological power? Furthermore, is the refutatory use of thought experiments itself justified?

cases” or “cases.” I will follow this custom and use these names interchangeably.

³ Cappelen and Deutsch have argued against this view by claiming that thought experiments are used mainly as counterexamples and intuition does not have any justificatory or evidential role to play (Cappelen 2012; 2022; Deutsch 2009; 2010). For a more comprehensive account of their arguments, together with responses against their views, see Climenhaga (2018).

These are the questions that I want to address in the present paper. In order to do so, I will commence in §2 by showing a specific feature of philosophical theories and showing how thought experiments relate to this feature. It is in this section where I answer the first question. In the following §3 and §4 I comment upon the works of experimental philosophers and deduce some troublesome consequences for the use of thought experiments. These two sections set the stage for §5, where I lay out three arguments to argue against the use of thought experiments as argumentative devices. Finally, in §6 I round off by making some final remarks.

§ 2. The Modal Status of Philosophical Theories

Just *a* case is enough to refute a well-established philosophical theory. This occurs due to a noteworthy feature of philosophical theories and analyses that is rarely noticed and spelled out:⁴ their modal status. Precisely, in order to comprehend the immense power that thought experiments have as refutation devices, it has to be acknowledged that philosophers present their theories presupposing that they should hold in every possible world, *i.e.*, to be necessarily true. Hence, the refutatory use of thought experiments presupposes the following principle:

A philosophical theory is true if and only if it is necessarily (\Box) so.

This principle allows to account for the fact that a mere far-fetched possibility (\Diamond) can refute an entire well-established philosophical theory. Furthermore, it shows us the usual argumentative pattern that thought experiments follow: “someone assert[s] $\Box p$, and an interlocutor rejoin[s] “but wait; $\Diamond q$, and $\Diamond q \rightarrow \Diamond \neg p$; therefore it is not the case that $\Box p$ ” (Hales 2009, 22).

In order to fully grasp this, let us consider Jackson’s thought experiment “Mary the neuroscientist.” Why does this case work as a valid argument against physicalism? Because it points to a possible scenario where it seems that Mary, although she possesses all the relevant material/physical knowledge, still discovers something new when she experiences colours; but if physicalism were true, then this should not happen. We should not think that she discovers something new. Thus, it is because it is assumed that physicalism is a necessary true theory which accounts for the fact that Mary’s case is sufficient to overthrow physicalism.

The natural world is usually regarded as the *de facto* place for scientists. With some exceptions here and there, it is usually the case that philosophers are neither trained to deal with the subtleties of experience, nor they tend to present their theories in a way for these to be verified against the tribunal of experience. This

⁴ With the following exceptions: Cohnitz (2006) and Kung (2016; 2021).

might explain the success story of thought experiments; they are extremely well suited for the purpose of checking the modal credentials of philosophical theories due to one of its fundamental features: the case needs only to be imagined. To be sure, in order to rebut the idea that meaning is in the head, we just need to imagine a twin earth where water's composition is XYZ instead of H₂O. An imaginative act of reason alone serves to confute a theory, or, as Timothy Williamson has put it, "much of the philosophical community allows that a judicious act of the imagination can refute a previously well-supported theory" (2007, 179).

As it transpires from the above, imagination plays a key role in the refutative use of thought experiments. Moreover, if they are to have such a destructive effect –as they actually do– it is because we are tacitly accepting the idea that imagination (or conceivability) is a reliable epistemic guide to possibility. We have to remark, then, that this in turn presupposes the endorsement of a conceivability-based account of the epistemic value of imagination.⁵ At their core, these accounts maintain that if we can imagine p , then $\diamond p$ and, on the contrary, if we cannot imagine p , then $\neg\diamond p$. For instance: since I can imagine turtles walking really fast, I can safely conclude that this is possible. Nevertheless, since I cannot imagine a squared triangle I thereby conclude that this is not possible.

The last example indicates that there are limits to what can be imagined, and, more importantly, there must be such limits, otherwise by imagining *n'importe quoi* we could prove or disprove theories indiscriminately. If someone can imagine a possible world where the logical positivists managed to verify the principle of verification, then it follows that it is possible to verify the principle of verification. I can imagine myself one morning opening a philosophical journal which claims with irrefutable certainty that the principle has indeed been verified. After imagining myself going through the journal, I imagine myself turning on the TV and seeing on the news channel "BREAKING NEWS: The principle of verification has been verified!" Now, would this conceivability entail actual possibility? It would not, since the very semantical principles that the principle conveys excludes itself from verification. It cannot thus verify itself; it is logically impossible to do so.

This example shows that imagination provides a way in which to constrain imagination in order for it to work as an effective guide for modal possibility. I contend that imagining x cannot provide a reliable epistemic role if (1) imagining x

⁵ Although there has been some critiques to conceivability-based accounts (Van Inwagen 1998), these theories (as expounded by Yablo 1993; Chalmers 2002; Kung 2010) seem to enjoy contemporary endorsement (Lam 2021). Furthermore, certain contemporary accounts do not require imagination to have imagery power. See, for example, Kung (2021) for an investigative account of the role of conceivability without imagery.

entails a contradiction and if (2) it is already known in the actual world that x is logically or semantically impossible. In having thus a constrained account, we assure an epistemological use for imagination (Kind 2018, 239; Peterson 2021, 227-228).

As a final point, it is important to mention that a thought experiment can only have argumentative force if and only if someone considers it to be a reason for or against a certain theory. If someone does not judge Searle's Chinese room as a reason against the computational theory of the mind, then the thought experiment by itself does nothing. This is due to the simple fact that if we cannot see how the thought experiment is pointing to a possibility where the theory does not hold necessarily, then we will not employ the thought experiment as a reason against it. Although it may seem trivial, from this point follows a crucial consequence: if someone – or even worse, an entire group of people –, when considering a thought experiment, does not see why the case would be a counterexample (a reason) against a theory, then it seems that the theory will still be intact. This corollary is not only an implication of the inner workings of the method of cases, but is rather an established fact documented by experimental philosophers.

§ 3. Experimental Philosophy, or From the Arm-chair to the Field

Metaphilosophical worries and issues regarding the use of thought experiments started to slowly emerge after the publication of Weinberg, Nichols and Stich's paper (2001). In this paper the authors investigated whether people from different cultural and linguistic groups would consider the same thought experiment differently. For verifying the hypothesis, they went on to conduct a series of experimental studies which confirmed the following two points: (1) Epistemic intuitions vary from culture to culture and (2) epistemic intuitions vary from one socioeconomic group to another. They corroborated these points by presenting to "Western" and "East-Asian" audiences a variation of a Gettier case and subsequently asking them whether the person in the scenario (A) really knows or (B) only believes. The result was "striking (...) a large majority of Westerners give the standard answer in the philosophical literature, viz., "Only Believes." But among Eastern Asians this pattern is actually *reversed*! A majority of EAs say that Bob really knows" (*op. cit.*, 443).

Sixteen years later, Machery and collaborators conducted similar experiments with people from 23 different countries to see whether "the Gettier intuition is robust across cultures and languages" (Machery *et al.* 2017, 532). Although the evidence that they found supports the claim, they nevertheless remark that "Bedouin data may be a counterexample to the claim that the Gettier intuition is universal" (530). Hence, there is at least a group of people who does not judge Gettier cases as reasons against the JTB analysis.

The field of work of experimental philosophers does not circumscribe to epistemology, as the above examples might indicate. On the contrary, their studies and findings extend well beyond this field.⁶ In the context of the present essay, the works of experimental philosophers on philosophy of language are of especial importance. Machery *et al.* (2004), for instance, investigated whether North-Americans' and East Asians' intuitions about reference support descriptivism or the causal-historical view of reference. For this aim they presented people from the United States and from Hong Kong variations of Kripke's Schmitt/Gödel and found that Kripke's cases elicit culturally variable intuitions: "Chinese participants tended to have descriptivist intuitions, while [US Americans] tended to have Kripkean ones" (*op. cit.*, B12). Similarly, Koch and Wiegmann (2020) have recently argued that, according to their experiments and findings, native English speakers' folk intuitions rather support what they call the "causal source view" of reference defended by Evans and Devitt.

Experimental philosophical studies show how cultural, linguistic and socio-economic variables, which are set aside in traditional philosophical reasoning, play an important role in making somebody judge a thought experiment as an effective reason for or against a philosophical theory. Moreover, there are also further factors that play such a role, like framing and order effects. These two features account for differences in judgement caused by variations in (A) irrelevant narrative factors and (B) the order in which cases are presented (Machery *et al.* 2018). Other variables that might affect our judgement of cases are gender, age, personality and academic affiliation (Stich & Tobia 2016). All this implies that a thought experiment is not judged as a reason solely due to its soundness or coherence. Rather, experimental studies make manifest the many different elements beyond the philosopher's control which influence the consideration of a thought experiment as an effective argumentative device.

Responding against experimental philosophy, arm-chair philosophers have argued that the findings of experimentalists are of no particular value for constructing or refuting philosophical theories for the same reasons that laypeople's beliefs about physics or biology do not count as reasons against physical or biological theories. After all, why should philosophers be interested in what laypeople think about philosophical cases? Philosophers should rather rely on their own expert consideration of thought experiments. This is why, according to Ludwig, "using surveys of untrained people to settle issues where there are conceptual knots in our thinking is fundamentally misguided" (2007, 149). Despite the initial plausibility of

⁶ For a comprehensive survey of the many areas in which experimental philosophers have worked and their results, see Sytsma and Buckwalter (2016).

this line of defence, the available empirical findings sustain the claim that even philosophers themselves, when thinking and considering thought experiments, are subjected to non-philosophical factors that influence their consideration of cases (Horvath & Koch 2021). Hence, despite Deutsch, who writes that “Gettier refuted the JTB theory, if he did, and Kripke refuted descriptivism, if he did, by presenting counterexamples, full stop. Whether these counterexamples are intuitive for anyone is a separate, and purely psychological, matter” (2010, 448), it has to be acknowledged that there are indeed different factors outside the philosopher’s control which do contribute to the way in which someone judges a thought experiment.

§ 4. Troublesome Consequences

I want now to draw three general points from what I have written so far: (A) philosophical theories assume a necessary modal status; (B) thought experiments, if they are to work as argumentative devices, depend on someone judging the case as an actual reason for or against a theory, and (C) there are extra-philosophical variables such as cognitive make-up and socialisation which influence the consideration of thought experiments.

These three points taken together have troublesome consequences for the argumentative use of thought experiments. We can see how by considering that one reason that we might have for thinking that internalism is incorrect is a thought experiment. But what if a group of people, among which there are philosophers, does not consider the thought experiment as a counterexample to internalism? What does it follow from here? Does it follow that internalism is true for them but false for us? How do we know who is right? Before answering these questions it is important to understand where exactly the problem resides. According to Stephen Stich:

Theories (...) assume that the contents of intuitive judgments are likely to be true. But if one group of people have the intuition that the protagonist in a thought experiment knows that p (or that her action was morally wrong), and another group of people have the intuition that the protagonist does not know that p (or that her action was not morally wrong), then obviously these two groups cannot both be right. So, unless the philosopher who is using intuitions as evidence for an objective phenomena theory can give a plausible reason why the intuitions of one group (typically the group that disagrees with him!) can be ignored, *demographic differences pose a fundamental challenge to the venerable philosophical tradition of using intuitions as evidence for objective phenomena theories* (2018, 385. Italics in original).

As I read him, Stich suggests that the problem that emerges when two groups have diverging intuitions resides in the fact that philosophers assume that the

content of the intuitive judgement elicited through a thought experiment is likely to be true. Hence, when faced with two different responses to the same case, a manner to select between the two would consist in (1) providing a meta-criterion for deciding between the two competing intuitions or (2) putting forward reasons that would explain away the intuition of one group over the other. I believe that both answers have fundamental problems. First, notice that (1) and (2) would be *ad hoc* strategies. In effect: philosopher *p* has found that a group of people does not see their thought experiment as “intuitive” (in Stich’s wording) and *then* presents reasons or a meta-criterion to explain away the conflicting intuition. But – and this is a second difficulty – even if they do this, on what grounds would they do it? How would they ground their arguments or the criterion? They might base them on extra-philosophical factors that bear on the discussion but cannot be philosophically settled or in further thought experiments, since these are taken to be the very foundations from which philosophical arguments gain their appeal. Therefore, philosopher *p* would need to go outside of philosophy or move in a circle. But be that as it may, one thing is clear: there is no apparent manner of solving the problem *within* philosophy. To my judgement, Stich’s way of articulating and solving the problem is unsatisfactory, being that it cannot be resolved within the confines of philosophical theorising.

But then, where does the problem lie? I believe that it lies in the argumentative use of thought experiments. For if thought experiments are supposed to be the fundamental reasons from which philosophical arguments and theories gain their appeal and if these reasons, in turn, are supposed to provide incorrigible or infallible justification, then it is because of the very argumentative use of thought experiments the ground that explains why the problem emerges in the first place. Indeed, were one not to use thought experiments as the base of one’s own theory or use them to criticise others theories, then there would not be any issue in having conflicting reactions to the same case. Furthermore, the argumentative use of thought experiments poses a radical metaphilosophical problem. For if two reactions to the same thought experiment cannot compete, then this very usage of cases eventually brings philosophy to an impasse. Due to these considerations, in the next section I will argue that the use of thought experiments for argumentative purposes should be abandoned.

§ 5. The End of the Case

In this section I will present the arguments that will justify what I have written above. Before proceeding to do so, I want to state the core idea that structures the arguments, to wit: since as a matter of fact there exist differences in the cognitive

constitution and socialisation of people, and since this constitution is partly responsible for the variations in their responses to thought experiments, then this variation can be exacerbated by going to the modal domain – which is not and needs not be encircled by real facts or the laws governing those facts⁷ – so as to imagine a possible world where people only exhibit differences in their cognitive make-up and thereby always judge thought experiments differently. In other words, the idea consists in imagining a possible world in which, for every philosophical theory $p_1, p_2 \dots p_n$, there is a group of people $g_1, g_2 \dots g_n$ who always regards a thought experiment as a reason against $p_1, p_2 \dots p_n$, pointing hence to situations where the theory does not hold necessarily.

This idea has a further positive consequence: experimental philosophers have established, using non-aprioristic methods, that people do not judge a determined thought experiment in the same way. For some, a Gettier case is a scenario where one possesses a justified true belief but not knowledge, whereas for others it is indeed knowledge.⁸ Nevertheless, one could circumvent the results of these studies by disputing the data, adhering to an *ad-hoc* hypothesis or claiming that philosophy's business has nothing to do with empirical findings. My arguments, however, since they are *a priori* and are presented in the modal domain, cannot be dismissed so easily by arm-chair philosophers.

As a result of this, and as I will later show, the value of thought experiments as a philosophical methodology is called into question. For what is the point of using a method which will constantly refute theories? From here then two alternatives are possible: either we abandon the modal status of philosophical theories or we discard the use of thought experiments for argumentative purposes. Both points have profound metaphilosophical consequences.

⁷ Cooper (2005) would disagree. She writes “we can say that a thought experiment is more likely to succeed if the thought experimenter is knowledgeable about the relevant aspects of the actual world. Only if she possesses either explicit or implicit knowledge of the behaviour of real phenomena can the thought experimenter predict how hypothetical events would unfold” (343). I do not agree with this requisite. Think about Thomson's violinist thought experiment (1971). Can one *really* attach a virtuoso violinist to the body of human beings? No, not really. Nevertheless, Thomson's thought experiment functions as a reason for defending the permissibility of abortion. To my understanding, this shows that philosophers do not need to have explicit or implicit knowledge about the behaviour of real phenomena. The same could be said about Davidson's swampman thought experiment (1987).

⁸ Is it then that they just do not get it? Are these two groups talking past each other? If it were only a case, then I would think so. But experimentalists' results have repeatedly shown that people disagree over many different cases. It is therefore a simple way out to maintain, *ad hoc*, that they do not understand the case. Hence, the evidence gathered so far is better explained by claiming that they do get it, but they just do not see the case as the philosopher wants them to.

With that said, it is now time for the arguments. Here is the first one, which is structured as follows:

Argument 1:

- (1) If thought experiments are a valuable metaphilosophical methodology, then by using them we can justify or refute philosophical theories.
- (2) If a thought experiment is to be used for justifying or refuting a theory, then there cannot be an impasse in people's consideration of them.
- (3) There are indeed impasses due to variation in people's judgement of thought experiments. The variation can be either *factual* or *modal*.
- (4) Hence, by (3) and (2) it follows that thought experiments cannot be used to justify or refute a theory.
- (5) Therefore, by (4) and (1) it follows it is not the case that thought experiments are a valuable metaphilosophical methodology.

Although I take the argument to be straightforward, I deem it best to explain the premises and the conclusions: Premise (1) is a conditional that establishes that if the method of cases is a valuable methodology, this is, a truth-conducive method, then philosophers can resort to this method in order to justify or refute a philosophical theory. Conditional (2) says that if a philosopher is to use a thought experiment for justification purposes (*i.e.*, for justifying or refuting a philosophical theory), then there cannot be stand-offs between two matching "intuitions". Differently said: if p_1 judges thought experiment TE as a reason against a theory and p_2 does not, then both p_1 and p_2 need to modify their judgements about the case if the thought experiment is to function as an actual reason for or against the theory; otherwise TE is used by p_1 and p_2 as a reason and not as a reason against the same theory. Now, since the judgements of p_1 and p_2 are incorrigible and infallible, it follows that p_1 and p_2 cannot come to modify their judgements. Hence, a stalemate is reached. Premise (3) establishes that there is indeed variation. This is the crucial premise of the argument, and in the following paragraphs I will justify it accordingly. Finally, conclusions (4) and (5) draw the consequences.

Here is a variation of the first argument:

Argument 2:

- (1) If the method of cases is a sound metaphilosophical method, then by using it we should not refute theories *ad infinitum*.
- (2) Using thought experiments we can indeed refute philosophical theories *ad infinitum*.
- (3) It is not true that the method of cases is a sound metaphilosophical method.

I take this second argument to be more direct than the previous one and in no need of further elucidation. Only a minimal remark is in place: I write

“metaphilosophical method” because here I am just restricting myself to criticise the philosophical use of thought experiments. In science they are used in a different manner and yield different results. Hence the importance of clarifying that I am here only addressing the philosophical use of it.⁹

In what follows I will present two sub-arguments and a thought experiment, which will justify premise (3) of the first argument. In due time, I will make clear what arguments justify premise (2) of the second argument.

The first argument, the “argument from variation,” consists in putting to work the evidence amassed by experimental philosophers and concluding that there is factual variation in people’s responses to cases.

Argument from variation:

- (1) If there is documented variation in people’s judgement of thought experiments, then this constitutes evidence for the factual variation of people’s judgements to thought experiments.
- (2) There is evidence registering people’s variation to thought experiments.
- (3) Therefore, there is factual variation in people’s consideration of thought experiments.

This sort of argument is usually controversial in the metaphilosophical debate between arm-chair and experimental philosophers for two main reasons. First, because it is disputed the amount and exact type of variation that actually exists and how this is problematic for the method of cases; second because there are not uncontroversial beliefs on what exactly the empirical data shows (Suhler 2019). For these reasons, I deem low the chances of success of this argument. Now, despite the fact that the argument can be thus challenged, it is already showing that is documented variation in people’s responses to thought experiments.

Here is the second argument (which also works as a justification of premise (2) of the second argument):

Imaginative power’s argument:

- (1) If it is conceivable to devise a scenario in which there is modal variation in people’s judgements of thought experiments, then this constitutes evidence for the modal variation of people’s judgements to thought experiments.
- (2) It is indeed conceivable to devise such a scenario.
- (3) Therefore, there is modal variation in people’s judgement of thought experiments.

⁹ See Schindler and Saint-Germier (2021). The authors critically analyse the role of thoughts experiments in physics and argue that “there is no ground for thinking that the method of cases is a somehow intrinsically flawed methodological oddity” (25).

Let me break down the argument and analyse it. Premise (1) establishes that if it is possible, in principle, to construct a thought experiment in which people react differently to a thought experiment, then this would constitute *prima facie* corroboration of the consequent. The method of cases, were it a good method, should not allow us to construct imagined scenarios where people diverge in their consideration of thought experiments. Premise (2) maintains that it is possible to imagine such a case; and here is where the exacerbation spoken of at the beginning of this section comes in to play a role. For surely it is more than possible to imagine such a case, if there already exists such variability as a matter of fact. Put it differently: if already constricted by the limits of the real world we find that some people do not see why a case should count as a reason against a theory, then this fact can be exacerbated to a much greater extent by the inner workings of the method of cases, for the sole faculty of imagination is enough for having twin earths, brains in bats, teleportation devices, zombies, experience machines and so forth. Hence, consequence (3). Unlike the previous argument from variation, this argument does not rely on the findings of experimental philosophers but rather on the inner workings of the method of cases: its imaginative dimension.

In the face of the previous point, a pressing question emerges: can we really imagine such a case? I contend that we can: imagine that in the future there lives a philosopher who, making the most of the technology of their days, invents a computer program that devises counterexample thought experiments for every philosophical theory proposed so far. The program does this by analysing the modal structure of the theories and then devising possible cases in which the theory does not hold, refuting it thereby. The program, being run by a supercomputer, invents the required case in matters of seconds¹⁰ and also indicates the required cognitive, psychological and social make-up that someone would need to possess in order to see the case as a counterexample. Now, since the case by itself will not refute the theory – somebody needs to regard it as an effective reason/counterexample –, the philosopher invents a second program for designing and making human beings (similar to what the game *The Sims* does). Using it, the philosopher designs a human being with the required cognitive, social and psychological features needed for ensuring that this person sees the case devised by the first program as an effective reason against the theory. Next, capitalising on the state of the art of cloning and human-design software and machinery, the philosopher proceeds to create the envisaged human. Once this is done, the philosopher asks them whether they think that the theory holds in the scenario devised by the first programme and they

¹⁰ The supercomputer in this thought experiment should be thought of as a maximiser of our own cognitive faculties. See Priest (2021).

answers with a simple “no:” they do not think that the theory holds in the possible world devised by the first supercomputer. Next, the philosopher repeats the process, but this time designs a human with the cognitive, psychological and social make-up required to make them see the case as a situation in which the theory does hold. In this way, here is an imagined situation in which there is modal variation in people’s judgement of thought experiments.

There is a further consequence to be derived from this thought experiment. Philosophers, were they still using thought experiments for truth-conducive purposes, could neither attain true philosophical theories nor could they resort to thought experiments to refute different theories. For each of these two purposes the philosopher of the thought experiment would respond appropriately. For the first case they would create a counterexample, and for the second they would design someone who will not consider the proposed case as a counterexample. Hence, in this futuristic world the use of the method of cases would imply the refutation of theories *ad infinitum* and the argumentative futility of thought experiments..

§ 6. Conclusion

In this paper my main objective has been twofold: to account for the fact that usually just a thought experiment is enough to undermine an established theory and to criticise a fundamental metaphilosophical tool used in various debates in philosophy.

I have shown that the destructive power of thought experiments resides in the modal status of philosophical theories, while the arguments and the thought experiment exposed in the previous section have justified the second claim. Furthermore, the arguments laid out in the previous section show that any argumentative use of thought experiments, positive or negative, is not truth-conducive. This implies that the conjunction of (A) the use of the method of cases for truth-conducive purposes and (B) the idea that philosophy deals with necessary true theories cannot be both maintained at the same time. In principle, we could abandon one of these constituents and embrace either one of these two views: to carry on with the idea that philosophers’ business is to construct necessary true theories and analyses, at the cost of abandoning the use of the method of cases for truth-conducive purposes. Alternatively, we can discard the idea that philosophical theories should be necessarily true and continue to use the method of cases. We could conceive the objective of philosophy to be that of investigating the concepts relative to a language, culture or background. Be that as it may, one thing stands fast: the adoption and implementation of either of these two standpoints would carry

Santiago A. Vrech

with it a reconfiguration of philosophy's conception or methodology. I hope to have shown that we have reasons to do so.¹¹

References

- Andow, James. 2020. "Intuitions about cases as evidence (for how we should think)." *Inquiry*: 1–33.
- Cappelen, Herman. 2012. *Philosophy Without Intuitions*. Oxford: Oxford University Press.
- . 2022. "Experimental Philosophy without Intuitions: An Illustration of why it fails," *Philosophical Studies* 179: 309–317.
- Chalmers, David. 2002. "Does Conceivability Entail Possibility?" In *Conceivability and Possibility*, edited by T. S. Gendler and J. Hawthorne, 145–200. Oxford: Oxford University Press
- Climenhaga, Nevin. 2018. "Intuitions are used as evidence in philosophy," *Mind* 127 (505): 69–104.
- Cohnitz, Daniel. 2006. *Gedankenexperimente in der Philosophie*. Paderborn: Mentis.
- . 2020. "Thought Experiments and the (Ir-)relevance of Intuitions in Philosophy". In *Philosophy in the Age of Science?: Inquiries into Philosophical Progress, Method, and Societal Relevance*, edited by Julia Hermann, Jeroen Hopster, Wouter Kalf, and Michael Klenk, 95–110. London: Rowman & Littlefield.
- Cooper, Rachel. 2005. "Thought experiments." *Metaphilosophy* 36 (3): 328–347.
- Davidson, Donald. "Knowing One's Own Mind." *Proceedings and Addresses of the American Philosophical Association* 60 (3): 441–458.
- Deutsch, Max. 2009. "Experimental philosophy and the theory of reference." *Mind & Language* 24 (4): 445–466
- . 2010. "Intuitions, counter-examples, and experimental philosophy." *Review of Philosophy and Psychology* 1 (3): 447–460.
- Hales, Steven. 2009. *Relativism and the Foundations of Philosophy*. Cambridge: MIT Press.
- Horvath, Joachim, and Steffen Koch. 2021. "Experimental philosophy and the method of cases." *Philosophy Compass* 16 (1): 1–13.

¹¹ **Special acknowledgements:** I would like to thank Annemarie Kalis, Daniel Cohnitz, Jesse Mulder & Bjørn Jespersen for valuable comments on earlier versions of this paper.

- Kind, Amy. 2018. "How imagination gives rise to knowledge." In *Perceptual Memory and Perceptual Imagination*, edited by Fiona Macpherson and Fabian Dorsch, 227–246. Oxford: Oxford University Press.
- Koch, Steffen and Alex Wiegmann. 2020. "Folk intuitions about reference change and the causal theory of reference." *Ergo: An Open Access Journal of Philosophy*.
- Kung, Peter. 2010. "Imagining as a guide to possibility." *Philosophy and Phenomenological Research* 81 (3): 620–663.
- . 2016. "Thought Experiments in Ethics." In *Knowledge Through Imagination*, edited by Amy Kind and Peter Kung, 227–246. Oxford: Oxford University Press.
- . 2021. "Why We Need Something Like Imagery." In *Epistemic Uses of Imagination*, edited by Christopher Badura and Amy Kind, 23–43. New York: Routledge.
- Lam, Derek. 2021. "An Imaginative Person's Guide to Objective Modality." In *Epistemic Uses of Imagination*, edited by Christopher Badura and Amy Kind, 44–62. New York: Routledge.
- Ludwig, Kirk. 2007. "The Epistemology of Thought Experiments: First person versus third person approaches." *Midwest studies in philosophy* 31: 128–159.
- Machery, Edouard *et al.* 2017. "The Gettier Intuition from South America to Asia." *Journal of Indian Council of Philosophical Research* 34 (3): 517–541.
- Machery, Edouard, *et al.* 2018. "Gettier was framed." In *Epistemology for the Rest of the World*, edited by Stephen Stich, Masaharu Mizumoto, and Eric McCready, 123–148. Oxford: Oxford University Press.
- Machery, Edouard, Ron Mallon, Shaun Nichols, and Stephen Stich. 2004. "Semantics, cross-cultural style." *Cognition* 92 (3): B1–B12.
- Peterson, Eric. 2021. "Attention to Details: Imagination, Attention, and Epistemic Significance." In *Epistemic Uses of Imagination*, edited by Christopher Badura and Amy Kind, 218–234. New York: Routledge.
- Priest, Graham. 2021. "Metaphysical necessity: A skeptical perspective." *Synthese* 198 (8): 1873–1885.
- Schindler, Samuel and Saint-Germier, Pierre. (2021) "Armchair physics and the method of cases." *Inquiry*: 1–25.
- Shoemaker, Sydney. 1994. "Lecture II: The Broad Perceptual Model--Self Knowledge and 'Inner Sense'." *Philosophy and Phenomenological Research* 54 (2): 271–290.

Santiago A. Vrech

- Stich, Stephen and Kevin Tobia. 2016. "Experimental philosophy and the philosophical tradition." In *A companion to experimental philosophy*, edited by Justin Sytsma and Wesley Buckwalter, 5–21. Malden: Wiley Blackwell.
- Stich, Stephen. 2018. "Knowledge, intuition, and culture." In *Metacognitive Diversity: An Interdisciplinary Approach*, edited by Joëlle Proust and Martin Fortier, 381–394. Oxford: Oxford University Press.
- Suhler, Christopher. 2019. "Why the method of cases doesn't work." *Review of Philosophy and Psychology* 10 (4): 825–847.
- Sytsma, Justin and Wesley Buckwalter. 2016. *A companion to experimental philosophy*. Malden: Wiley-Blackwell.
- Thomson, Judith Jarvis. 1971. "A Defense of Abortion." *Philosophy and Public Affairs* 1 (1): 47–66.
- Van Inwagen, Peter. 1998. "Modal epistemology." *Philosophical Studies* 92 (1/2): 67–84.
- Weinberg, Jonathan, Shaun Nichols, and Stephen Stich. 2001. "Normativity and epistemic intuitions." *Philosophical topics* 29 (1/2): 429–460.
- Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.
- Yablo, Stephen, 1993. "Is conceivability a guide to possibility?" *Philosophy and Phenomenological Research* 53 (1): 1–42.