

THE END OF THOUGHT EXPERIMENTS?

Mark MALLER

ABSTRACT: This reply is a refutation of Santiago Vrech's article "The End of the Case? A Metaphilosophical Critique of Thought Experiments" (2022) which argues that thought experiments used in argumentation cannot hold in All Possible Worlds (APW) modality, and thus should end. Cases are used to justify or refute a philosophical theory, but should not have the power to refute an entire theory, especially *ad infinitum*. Significant variations in intuitions, he argues, invalidate cases and are not proven. I argue some variation is acceptable and expected, and his bar of criteria is too high—APW validity is unnecessary. Vrech's three arguments are faulty and/ or too weak, and his thought experiment is of questionable value. Experimental and armchair philosophy (or philosophers) are misnomers, misleading and false dichotomy.

KEYWORDS: thought experiments, metaphilosophy, logical argumentation, All Possible Worlds, modal logic

Contrary to Vrech's article "The End of the Case? A Metaphilosophical Critique of Thought Experiments" (2022), cases are alive and well. He attempts to disprove their validity as argumentative devices (strategies), particularly in the modal All Possible Worlds (*APW*), and uses three arguments in order to show that a philosophical theory must hold in *APW* to have validity. This counter-reply analyzes these arguments, and shows that his attempt is illogical and unjustified. I raise objections to Vrech's theory and show advantages of thought experiments, and criticize the distinction between experimental and armchair philosophers. Thought experiments (cases), are used either to defend a position, or to refute a theory through counter-examples. I argue that discontinuing them is an error and disservice to philosophy. Vrech rightly argues they ought not have the power to refute or overthrow an entire theory; certainly not highly established ones like physicalism or free-will. He is not opposed to their use in cases intended for provoking thought, for questioning ideas, and to ponder possible answers to open questions, yet most cases serve both purposes.

Thought Experiments: for and against

Vrech defends his point with Frank Jackson's thought experiment "There's Something about Mary" (1982), which attempted to refute physicalism by showing that Mary, the neuroscientist, possesses complete scientific knowledge of color but

Mark Maller

no experiences, and lives in a black and white world.¹ Mary's new experience of seeing all colors intends to prove that her factual knowledge is incomplete, therefore, physicalism is wrong. I agree that this case alone should not (and does not in the long-term) have the power to refute physicalism. The APW logical modality is seemingly not utilized here, and Jackson changed his mind years later. Similarly, other cases which intend to refute entire theories and rely on data in neurophysiology would be suspect.²

Regarding the power of thought experiments, Vrech argues—

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 so. (163)

Now this sweeping generalization for all possible worlds is untrue for a multitude of thought experiments, (including my cases), which were never intended for APW. Like others, I do not presuppose that the case holds in every possible world, but at the minimum it holds in this world and that is sufficient. Utilizing the parameters of modal logic is unnecessary, over ambitious and conjecture. It is highly implausible that thought experiments will hold in APW because of the totality of complex counterfactuals from individual imaginations, and whether it cannot hold in one or more worlds is speculation and an insufficient reason to invalidate it. Thus, Vrech sets the criteria bar too high, which would deter or discontinue such future thought experiments. According to his strict modal logic, it may seem reasonable to those who abide by, or are committed to this David Lewis modality, and rarely utilize cases. To my knowledge, thought experiments neither imply or infer 'immense power' by themselves, which should refute a long-held theory, unless that theory is illogical or flawed for other reasons. Clearly these cases potentially can and do refute such theories, but this does not imply they *ought* to be refuted. Moreover, it is problematic that thought experiments are (or should be) remotely 'incorrigible or infallible justifications' for theories. In this metaphysical context, 'should' (ought) is used normatively with underlying implications. Those who defend this position

¹ Actually, black and white are real colors.

² Jackson's newer view is included in a collection of essays with the identical name (2004). In the foreseeable future, I predict physicalism will be proven right, though evidence is lacking presently. In 50-100 years scientists should be ready to prove that brain activity and physiology can explain all but rare human experiences, considering the past 75 years of neuro-anatomical progress.

are too sanguine and their counterfactual imaginations are extreme.³ Vrech rightly claims that conceivability in the imagination should have limits—otherwise examples may be pure fantasy, strain naïve credibility and often absurd. We can imagine fast turtles but not square triangles. We may imagine the verification principle is reportedly verified on the evening news, he quips. Searle's Chinese Room experiment is a strong argument, but philosophers do not hold it is true in *APW* or that it provides epistemic proof of his position. Slight disagreement in thought experiments should be deemed acceptable in that unanimity in philosophy is rare, and is a major reason why the field seems to makes no progress—partly because of innumerable minor differences or distinctions in theories. Expecting thought experiments to be true in *APW* is like requiring minor civil law cases to be tried with the same strict standards as criminal cases. In civil cases a verdict of guilt does not demand 100% proof. Lottery players cannot expect to win in any possible outcome, regardless of the number of tickets purchased. More examples of powerful cases would have strengthened his argument.

Thought experiments rely on intuitions from philosophers, and experimentally with the public. While these intuitions may vary widely with disagreement, they are one of the best available strategies in academic discourse. Other criteria include strong scientific evidence, logical validity, and common sense. Experimental philosophers contend philosophers hold no privileged status on wise intuitions. He notes some thought experiment studies have shown significant variations from different cultures, particularly Asians from Hong Kong, among other locations, and experimental philosophers claim that this difference is important. The belief is that Others possess astute intuitions, and different cultures reveal conflicting answers to thought-provoking questions.

The Intuitions of Others

I will argue that only a minority of educated agents possess intelligent intuitive skills for thought experiments. Evidence from eastern cultures may reveal interesting variations, but it is mixed and debatable, possibly with language concerns, such as ambiguity and translations of local slang and abstract concepts. Outsiders and students are unfamiliar with academic terminology, e.g., causation, empirical, sensory, determined. My surveys in Ethics and Introduction classes over many years

³ One prominent example is Alvin Plantinga's case for his defense of God, based on modal logic and All Possible Worlds. He attempts to prove that the omnipotent God must exist in *APW*, which must include a moderate amount of evil that is necessary for human freedom. However, criticisms show that it is possible that God can exist in some worlds, but it is highly implausible that He exists in *APW*, and perhaps not in this actual one.

Mark Maller

showed the opposite; that is, no significant differences among adult and traditional students from foreign backgrounds. Furthermore, statistical methodology and measuring the answers would be difficult or challenging, considering all the variables: age, gender, race, nationality. Vigeant (2024) questions whether Others should be involved.

The common advice of turning to the Other to assist with the evaluation and analysis of an argument is far from ideal. Individuals are likely to be flawed, in many of the same ways as the person developing the argument. Adding more individuals to the process will not necessarily help as they can amplify the original issues as well as introduce new ones. (53)

Diversity in itself is no virtue, she claims, in practice and in observations. Confirmation Bias of experimenters may influence studies, though this is not always a fault. The important question is whether their opinions should matter to philosophers. Although comparing philosophy cases to factual theories of biology and physics is irrelevant and weak analogy. Social scientists are aware of this challenge and adjustments are needed. The reasons why laypeople's opinions on cases have lower value are—1. Their attitudes are doubtful with no motivating reasons to take cases seriously. 2. Lack of critical thinking skills and imagination, insufficiently open-minded. 3. No familiarity with abstract ideas and terminology. Some subject knowledge is useful to judge and evaluate some thought experiments wisely. He quotes Cooper's insight.

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. (Cooper 2025 in Vrech 2022, 169)

He disagrees, using the example of Thomson's violinist experiment she created for the defense of permitting abortion. This notably weird case is unconvincing to me, though its conclusion is agreeable. Minimal knowledge of abortion would probably enable better thoughtful answers, and some men have only bare knowledge of pregnancy. Vrech seems unfamiliar with the wide variety of thought experiments in metaphysics, epistemology, and applied ethics in which knowledge of real life is important. A false assumption is that the public has knowledge of the same subjects as the interviewer because facts show that people are mostly uninformed, contrary to their belief.

The arguments, some rephrased or shortened.

Three Arguments and Thought Experiment

1. If thought experiments are a valuable methodology, then we can justify or refute philosophical theories.

2. If thought experiments are used for justification or refutations, then there cannot be an impasse on the agent's answers to them. *Weak*. Impasses would make the case difficult but not impossible to resolve. Temporary impasses could be deemed acceptable in some cases.

3. Impasses exist because of people's variations of judgments of thought experiments. Variations can be factual or modal.

4. Hence, by (3) and (2) it follows that thought experiments cannot be used to justify or refute a theory.

False. This leap to conclusion cannot follow. (2) is weak and (3) is partly true only. It is possible impasses are based on tentative or petty distinctions, and may be caused by other issues than factual or modal. Only some impasses are seemingly intractable. Two or more premises are necessary between (3) and (4). The argument needs reconstruction to be salvaged.

5. Conclusion. By (4) and (1) it follows that thought experiments are not a valuable methodology. Premises(s) are needed between (4) and (5). *False*. Fallacy of hasty generalization. *Unsound*.

His *second* argument is succinct and fares no better.

1. If the methodology of cases is sound, then we should not refute theories *ad infinitum*.

2. Using thought experiments we can refute philosophical theories *ad infinitum*.

3. Conclusion. "It is *not* true that the method of cases is a sound metaphilosophical method." (170). This brief argument needs elucidation, despite his disclaimer otherwise.

It is not proven that the methodology of cases is sound or unsound, or is a definite issue of validity, and this should be established first. It is uncertain whether all metaphilosophical cases can be categorized in one group or divided out by subject or kind, or how to determine the criteria for this project. To my knowledge, the problem of refuting or not refuting cases *ad infinitum* or *APW* is not an essential issue. Examples of these are lacking here. (3) is speculation, conjecture only. *Unsound*.

The *third* argument is a repetitive, valid modus ponens. *If A then B/. A/. Therefore, B*.

1. If there is documented variation in people's judgments of thought experiments, then this constitutes evidence it is factual.

Mark Maller

2. There is evidence registering variations to thought experiments.
3. Therefore, there is factual variation in people's consideration of thought experiments.

Although its significance is unproven, sufficient and anecdotal evidence exists for this tentative assumption. He acknowledges that the degree and type of variations are disputable, and the empirical data is questionable at present. Since he admits this argument will not be successful, why is it included? This wordy argument repeats what is already known or assumed, epistemically. His two versions of this *modus ponens* are weakly devised and do not strengthen the general argument. Thus, it is sound, but lacks sufficient force and interest to move or advance his argument along.

Ironically, Vrech defends his position with his own thought experiment. Rarely does an author argue against using an argumentative strategy, then utilize it himself unless as a trick. He presents a clever imaginative case to show "if it is possible in principle to construct a thought experiment in which people must necessarily react differently..." to the same experiment (172). He imagines a future when a philosopher invents a computer program that devises counterexample thought experiments for every philosophical theory proposed so far. The program analyzes the modality of the theories and devises possible cases whereby the theory does not hold, therefore, refuting it. Then the philosopher invents a *second* program which creates a being with a very human mind that ensures it sees the case of the first program as an effective reason against the theory. After analyzing the first program, the "person" believes the theory does *not* hold in the possible world. Alternately, the "person" can be redesigned to believe that the theory *holds* in the possible world. This contradictory answer, hypothetically, intends to show the methodology of cases is futile in this futuristic world and refutations would proceed *ad infinitum*.

Granted, this original thought experiment serves its purpose and holds interesting conjecture, like nearly everything in possible world theories. Some will accept the case for its purpose. Nevertheless, it cannot prove or add much to his argument that thought experiments are theoretical dead-ends and should be abandoned. It is his best argument, albeit ironic, and intentionally demonstrates a mere possibility, though it would be stronger as the first argument, not last. Of course, just as easily an imagined philosopher could prove the opposite: total agreement with the theory with no modal variation. In this alternate case, thought experiments and counter-example methodology become valid strategies.

Finally, there is hope for his suggestion. "We can discard the idea that philosophical theories should be necessarily true and continue to use the method of

cases.” (173) This is the best conclusion we can expect, but not that it implies a ‘reconfiguration of philosophy’s conception or methodology’. There is no need to change our conception of analytic philosophy; it is hugely broad, full of variety and kinds of logical argument to encompass all styles and methodology. Perhaps he is thinking of the narrow-focused work of language and logical-oriented theories of Kripke, Wittgenstein and the like—*not* philosophy of religion, Other minds, aesthetics, and self-identity problems. Finally, Vrech cannot show the modal use of thought experiments is *unconducive* to the truth, nor that using them to establish or refute a theory is logically futile. A pragmatic admonition is to analyze them with the proverbial grain of salt—perhaps a pound of sodium, and use greater caution toward granting or implying significant power to such cases because of fantastical ideas.

Moreover, thought experiments are often the preferred philosophical style, apart from the subject. They serve as illustrations—conceptual pictures within solid dense text. Metaphors bring a fresh breath in a desert of logical argumentation like a riddle. They can be the most interesting part of theories if used prudently with discretion, or they can inundate the argument with irrelevant examples and counter-examples that trivialize the position. Thought experiments and examples of taste sensations, low culture, humor and fantasy, reduce the seriousness of the argument and often weaken it, especially in ethics. The best thought experiments are remembered and/ or cited long after the argument itself is forgotten, like the famous Trolley case, now sold in stores. Thought experiments belong in philosophy as much as excellent examples. The history of philosophy is replete with influential cases, but unlike Occam’s Razor, they have multiplied beyond necessity. Plato created the first developed thought experiment, but the term was not used until the eighteenth century. The Gyges’ Ring story in the *Republic* is a convincing thought experiment.

Armchair and Experimental Philosophers

The putative difference between experimental philosophers and armchair philosophy is a false and misleading distinction, though Vrech accepts this difference in his article. Despite its potential promise, designing thought experiments exclusively is counter to philosophy’s traditional spirit and intention. Thought experiments are conceptual tools, not ends in themselves. Outsiders to a field do not grasp the issues and implications or only superficially, nor understand the concepts. (My students think that ‘causal’ means casual). Universally, agents tend to assume their knowledge is greater than it is, as experienced educators know. Egos exceed intelligence. One issue is whether the surveys are conducted in the field in person, or most likely, through the internet or in classes. If either of the latter two, then

Mark Maller

experimenters, too, are armchair philosophers, literally in office chairs. Philosophers are not trained as psychologists, and their methodology may or may not be scientifically valid or realistically scientific. For this reason, the distinction serves no important purpose. Why philosophers? We would not call Oppenheimer an armchair physicist, nor Turing an armchair mathematician. The public uses the depreciating 'Ivory Tower' to label individuals or ideas that are grossly unrealistic and divorced from life. The word 'armchair' is merely a step away. The word 'experiment' seems inappropriate because it connotes real labs, statistical testing or fieldwork and scientific projects. Perhaps the words 'Thought Exercise', 'survey or Thought Reflection' may substitute. The term 'armchair' has negative connotations of a soft lifestyle, disengagement with the actual world, and sounds disdainful. It should be abandoned.

Experimental philosophy is similar to social psychology in that the major purpose is descriptive reporting of intuitions and opinions of ordinary individuals. Conflating and confusing it with academic philosophical problems, typically, shows low significance or value, thus far. Arguably, experimental philosophy seems an oxymoron because it performs no real experiments nor utilizes them per se, except in interdisciplinary studies. Therefore, use of words 'experimental' and 'armchair' are misnomers in principle and practice. Scientific thought experiments of scientists are a different category which my argument does not apply.

Summary

This reply shows that thought experiments (cases) used to refute or defend a theory are useful and important in philosophical argument. Vrech's three arguments are faulty or too weak for a good defense, and his original thought experiment raises questions. Validity in All Possible Worlds is neither necessary nor sufficient, and thus ought not be the criterion for viability. The Bar of Criteria for refuting such arguments *ad infinitum* is too high. Variations in responses to cases surely exist, but this is not as problematic as he holds. Definite evidence is lacking at present but this is not necessarily a reason to deter using them.

References

- Clifford, W.C. "The Ethics of Belief". <https://people.brandeis.edu> (PDF). Originally 1877.
- Cooper, Rachel. 2005. "Thought Experiments". *Metaphilosophy* 36(3). 328-347. Quoted in Vrech 2022, 169.
- Jackson, Frank, 2004. "Epiphenomenal Qualia", "What Mary Didn't Know", Postscript and "Mind and Illusion". In *There's Something About Mary. Essays*

on Phenomenal Consciousness and Frank Jackson's Knowledge Argument, edited by Peter Ludlow, Yugin Nagasawa and Daniel Stoljar, 39-51, 51-57, 409-442. Cambridge: MIT Press.

Knobe, Joshua and Shaun Nichols. 2017. "Experimental Philosophy". In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. URL = <<https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy/>>.

Maller, Mark. 2022. "The Evil That Free Will Does: Plantinga's Dubious Defense". *Metaphysica* 23(1): 37-49.

Tittle, Peg. 2005. *What If? Collected Thought Experiments in Philosophy*. New York: Pearson. Includes Searles' "Chinese Room", Thomson's "Trolley and Violinist", Plato's "Gyges Ring", Jackson's color experiment of Mary. Over 120 cases.

Vigeant, Louise. 2024. "Generative AI and Argument Creativity". *Informal Logic*. 44(1): 44-64.

Vrech A., Santiago. 2022. "The End of the Case? A Metaphilosophical Critique of Thought Experiments". *Logos & Episteme. An International Journal of Epistemology*, XIII(2): 161-176.