

## **THE IMPORTANCE OF THOUGHT EXPERIMENTS**

**ABSTRACT:** Living in the world of legal norms seems to be easy: you have rules of conduct guiding you how to behave in a lot of life situations including the fact what will happen to you if you do not obey these rules. In a way, legal norms are predicting the future giving us the guidelines for living. Although the legal system together with its rules tend to cover all areas of social life, there are situations that couldn't be foreseen at the time of making a particular regulation. These gaps could be spanned by adopting subsequent rules of conduct. In order to predict an event that may occur, and to predict human behavior in these situations as well as a human response to punishment when someone violates a rule, it is good to conduct a thought experiment. The basis of a thought experiment can be a completely fictitious and even currently impossible event, or a variation of some of the known and possible events. The key question when formulating a thought experiment is "*what if*". The answers to this question may start with "then it is possible", "then it will be", "it could be" or something similar. The answers will differ in terms of content only on the basis of the values, beliefs and attitudes of the one who answers the "what if" question. In our paper, we will briefly present the concept of a thought experiment, its internal structure, types and, by giving some examples, encourage readers to be more informed about this topic.

**Key words:** *thought experiment, imagination*

---

\* Ph.D. Assistant Professor, The Faculty of Law, Novi Sad, The University of Novi Sad, Republic of Serbia, e-mail: d.coric@pf.uns.ac.rs

## 1. Introduction

Our scientific practice is rich of thought experiments, although their results may be subject to further empirical testing. A great part of ethics, philosophy of language, and philosophy of mind is based on the results of thought experiments in a way that seems very similar to scientific thought experiments (though some might contest this), including Searle's Chinese room, Putnam's twin earth, and Jackson's Mary, the colour scientist and many others (Brown & Fehige, 2019). Philosophy, more than other sciences is fully enriched with thought experiments and their results, that even some authors say that without them, a great portion of philosophy wouldn't exist. Legal practitioners use the thought experiments also a lot, especially when they are in need to establish new rules.

But what is the thought experiment?

A lot of definitions of thought experiments were made, or just tried to be made. All of them have the same, simplest idea that thought experiment is "*an experiment carried out in thoughts only*" (Merriam-Webster Dictionary, 2020, e-version), not in reality, in some laboratory or somewhere else outside. It is the use of an imagined scenario in order to help us understand the way things really are. The understanding of the situation and possible outcomes comes through reflection on the situation and thru pure process of thinking. "*Thought experiments are performed in the imagination*", says Brown (2019). "*We set up some situation, we observe what happens, then we try to draw appropriate conclusions*" (Brown & Fehige, 2019). They are "*what ifs*" drivers of our thinking about something. Yeates (2004) says that "*a thought experiment is a device with which one performs an intentional, structured process of intellectual deliberation in order to speculate, within a specifiable problem domain, about potential consequents (or antecedents) for a designated antecedent (or consequent)*" (p. 150). In short, during the thought experiments we gain new information by rearranging or reorganizing already known empirical data in a new way, looking it over from another perspective. In such a way, we stimulate our thinking, using our imagination, and we are testing our values, beliefs and standpoints at the same time.

Although the history of use of thought experiments is long and dates from Plato's cave or even earlier (Rescher, 1991, p. 31–41), it is considered that Hans Christian Ørsted was the first to use the German term *Gedankenexperiment* (lit. thought experiment) circa 1812 (Witt-Hansen, 1976, p. 48). Ørsted was also the first to use the equivalent term *Gedankenversuch* in 1820 (Witt-Hansen, 1976, 49). Ernst Mach (1960) used the term *Gedankenexperiment* in a different way, to denote exclusively the imaginary conduct of a real

experiment that would be subsequently performed as a real physical experiment by his students (pp. 32–41, pp. 159–62). After those two experiments, one that is conducted in reality and the same, conducted only in imagination, with words, not with actions, Mach interviewed students which consequences were more real to them, from the first or from the second experiment. The English term *thought experiment* was coined and first appeared in the 1897. in english translation of one of Mach's papers.

During the mid-1980s, thought experiments were recognized as a central technique in analytic philosophy. Thanks to the activities done especially by Tamara Horowitz and Gerald Massey, philosophers and other scientists started to think about thought experiments again. Few years of dedicated works resulted with conference at the University of Pittsburgh on "*The Place of Thought Experiments in Science and Philosophy*" and with collected works from this conference, published in 1991.

The most important work for popularization of thought experiments in 21<sup>st</sup> century is book "*What If...Collected Thought Experiments in Philosophy*", by Peg Tittle, which is published in 2016. It is a brief collection of over 100 classic and contemporary thought experiments, each exploring an important philosophical argument. All experiments are divided into 9 areas: metaphysics, philosophy of mind, personal identity, philosophy of language, epistemology, logic, ethics, social and political philosophy, and aesthetics. After every experiment, there is a possible solution or explanation of the author, which is not definitive, and it is given as an introduction to further thinking and solving mentioned situation. Everyone, no matter how much they know(about) philosophy, can give their opinion about the behavior of a third person in the context of that experiment or set themselves as the main character and ask themselves, *what if I was in that situation, what would I do*. That is, we can say, the beauty of thought experiments, especially in philosophy, because they can challenge everyone's opinion, at all times.

## 2. Types of thought experiments

According to Yeates(2004), there are seven types of thought experiments:

1. *Prefactual* thought experiments speculate on possible future outcomes (Sanna, 1998, pp.635–665) or *What will be the outcome if specific event happens*"(Yeates, 143). For example, "*what would happen to your friend, if you tell his wife everything you know about him*" (Yeates, 2004, p. 144). It could happen nothing, or something, in numerous ways.

2. *Counterfactual* thought experiments speculate on the possible outcomes of a different past, ie if past was different, how would it affect to someone particular (Goodman, 1947, p. 114). For example, what if I have chosen to study mathematics, would it be possible for me to be an university professor in that area?
3. *Semifactual* thought experiments speculate on the extent to which things might have remained the same, despite there being a different past (Goodman, 1947, p. 117). For example, imagine that it is raining and you have an umbrella but is it possible to get wet anyway and how?
4. *Prediction* in this case, we project the circumstances of the present into the future. Some authors say that scientific prediction takes two forms (Sarewitz & Pielke, 1999, p. 121):
  - (1) Using known principles of nature, we predict what might happen, or
  - (2) Using suites of observational data and sophisticated numerical models in an effort to foretell the behavior or evolution of complex phenomena.
5. *Hindcasting* involves running a forecast model after an event has happened, in order to test whether the model's simulation is valid. It is usually used for weather forecasting.
6. *Retrodiction* is moving backwards in time, step-by-step, in as many stages as are considered necessary, from the present into the speculated past to establish the ultimate cause of a specific event (e.g., reverse engineering and forensics). It means that "*past observations, events and data are used as evidence to infer the process(es) the produced them*" and that diagnosis "*involve[s] going from visible effects such as symptoms, signs and the like to their prior causes*" (Einhorn & Hogarth, 1982, p.115). Retrodiction is kind of post factum analysis and giving the opinion on something that is finished.
7. *Backcasting* (Yeates, 2004) is establishing the description of a very definite and very specific future situation. It then involves an imaginary moving backwards in time, step-by-step, in as many stages as are considered necessary, from the future to the present to reveal the mechanism through which that particular specified future could be attained from the present (p.158).

In literature can be found even more classifications of thought experiments, but we think that this one is the most accurate and the most comprehensive. The classification itself shows a wide application of thought experiments in various fields of life and science.

### 3. Inner characteristics of thought experiments

What differs thought experiments from other exclusively laboratory experiments? According to Letitia Meynell (2014), there are six inner characteristics:

1. *Imagery*, needed for stimulating the imagination and focusing on the very idea that is discussed. The line of imagination can be very fragile and unstable, so the one who is first imagining the experiment and conducting it must use:
2. *Experiential language*. He/she must behave as he/she did it in the reality, and have their own experience of what he/she are talking about. As the experiment-teller can more vividly and with a lot of details tell the imagined story, the more can he/she engage others to get involved and give their opinion.
3. *An epistemological analysis* is also rather important, thinks Meynell, because it shows how the thought experiment justifies (or fails to justify) its conclusion. It is a kind of paradox, that we expect that we can learn something new from the imaginary scenarios. But, if we clarify to ourselves that thought experiments and our solutions of the experiments are connected to our values, beliefs and attitudes, we can learn a lot –about ourselves and about others.
4. The thought experiments have “*the irreducibly imaginative character*”, says Meynell (p. 4151). Every solution of this imaginative problem is correct, to the one that has given it. That is why the thought experiments are considered as:
5. *Instruments of provocation and imagining different interpretations*. We learn even by provocation and investigation of something, not only by memorizing pure facts. Every person reflects his/her own values, beliefs and experience, while searching for solutions of the thought experiment.
6. Thought experiments are considered *as objective* even though they are not real( but they might be real). Everyone acts from the position of their standpoint, and for them, their solution is their objective truth.

Having in mind all these characteristics of thought experiments, we may say that they are true and very valuable companion in developing imagination and process of decision-making, in any area of life.

Some thought experiments have the scenario that *could be nomologically possible, or possible according to the laws of nature*. For example, John Searle's *Chinese room* is nomologically possible:

"Imagine a native English speaker who knows no Chinese. Locked in a room full of boxes of Chinese symbols (a data base) together with a book of instructions for manipulating the symbols (the program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions (the output). The program enables the person in the room to pass the Turing Test for understanding Chinese but he does not understand a word of Chinese.

The point of the argument is this: if the man in the room does not understand Chinese on the basis of implementing the appropriate program for understanding Chinese then neither does any other digital computer solely on that basis because no computer, qua computer, has anything the man does not have." (Searl, 1999, p. 214).

Searl's experiment, throughout decades, has showed that artificial intelligence has the eligibility to think and work without consciousness (when consciousness is defined as something characteristic only for humans) and what could be the results of it. Personal computers, as some authors say, have evolved from vigilant Eliza, and "*have moved from the lab to the pocket and the wrist*" (Cole, 2020). We are trying rather hard to give those artificial intelligences some consciousness and make them more alike to us. The only thing that is not clear is that we want them so similar to us, because we want to transfer all the burden of thinking and decision-making to them, or to have a neutral observer who will help us in life, at key moments.

Some thought experiments present scenarios that *are not nomologically possible but are metaphysical possible*. For example, Hillary Putnam's (1973) thought experiment is one of them:

We begin by supposing that elsewhere in the universe there is a planet exactly like Earth in virtually all aspects, which we refer to as "Twin Earth". (We should also suppose that the relevant surroundings are exactly the same as for Earth; it revolves around a star that appears to be exactly like our sun, and so on). On Twin Earth, there is a Twin equivalent of every person and thing here on Earth. The one difference between the two planets is that there is no water on Twin Earth. In its place there is a liquid that is superficially identical, but is chemically different, being composed not of H<sub>2</sub>O, but rather of some more complicated formula which we abbreviate as "XYZ". The Twin

Earthlings who refer to their language as “English” call XYZ “water”. Finally, we set the date of our thought experiment to be several centuries ago, when the residents of Earth and Twin Earth would have no means of knowing that the liquids they called “water” were H<sub>2</sub>O and XYZ respectively. The experience of people on Earth with water and that of those on Twin Earth with XYZ would be identical. Now the question arises: when an Earthling (or Oscar for simplicity’s sake) and his twin on Twin Earth say ‘water’ do they mean the same thing? (p. 700).

The result of this experiment is that the contents of a person’s brain are not sufficient to determine the reference of terms they use, as one must also examine the causal history that led to this individual acquiring the term. (Oscar, for instance, learned the word “water” in a world filled with H<sub>2</sub>O, whereas Twin Oscar learned “water” in a world filled with XYZ) .Meaning isn’t just in head, says Putnam, but also in experience of a person and his/her understanding of the circumstances in which he/she is , in his/her own environment.

In some cases, the hypothetical scenario might be considered *metaphysically impossible, or impossible in any sense at all*. Such thought experiment was conducted by David Chalmers (1996), regarding existence of zombies, or persons who are physically identical to us in every way, but who lack consciousness. Disputes about their conceivability goes for few decades and the result of these debates is fairly uniform as there are authors who think that the existence of a zombie as a person without any consciousness and conscience is entirely possible, while some authors firmly deny such a possibility. One thing is for sure, the Hollywood film industry supports the efforts of those who believe that zombies exist and that due to the physical deficiency of the brain, in which all functional centers are situated, we already have these zombies somewhere next to us (Kirk, 2019).

For thought experiment, there is also important their orientation in time (Yeates, 2004, 160). They can be either:

- *Antefactual speculations*: experiments that speculate about what might have happened prior to a specific, designated event. For example, Richard Sylvan’s (1973) experiment “*Last man argument*”, about possible behavior of the last man standing on Earth is an antefactual one experiment. Sylvan thinks about the behavior of the last man on earth, from different points of view of environmental ethics. If the last man on earth accepts an instrinsic value of everything, whether living or non-living, his last hours spent on earth will be focused on maintaining his life and the life and survival of other species. If, on the other hand, the last man accepts the anthropocentric opinion that only man

is worth living because he has the consciousness and power of thinking and planning, such a last man will activate the button for self-destruction of the whole planet, and all other plant and animal species, survivors of the cataclysm, because their survival is of no importance to him who is dying.

*-Postfactual speculations:* experiments that speculate about what may happen subsequent to (or consequent upon) a specific, designated event. Although it is not enlisted fully as thought experiment, but more as example of game theory, The Prisoners' Dilemma can be observed in this context:

“Two members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of communicating with the other. The prosecutors lack sufficient evidence to convict the pair on the principal charge, but they have enough to convict both on a lesser charge. Simultaneously, the prosecutors offer each prisoner a bargain. Each prisoner is given the opportunity either to betray the other by testifying that the other committed the crime, or to cooperate with the other by remaining silent. The possible outcomes are:

If A and B each betray the other, each of them serves two years in prison

If A betrays B but B remains silent, A will be set free and B will serve three years in prison

If A remains silent but B betrays A, A will serve three years in prison and B will be set free

If A and B both remain silent, both of them will serve only one year in prison (on the lesser charge)”( Kuhn, 2019).

The thought experiment can be here done by asking a question: *You are one of those two criminals, what would you do and what would you think that your best friend and companion in crime would do in this kind of situation?*

Also, we can say that there are thought experiments that are they past-oriented and future-oriented. Firstly, past oriented thought experiments are dealing with possible outcomes in the future if the initial events in the past were different. For example, If I would studied mathematics, would my life be different today? Or, like Ludwig Wittgenstein presumed in his thought experiment called “*The Big Book*” (Darwall, Gibbard, & Railton, 1997, p. 67) about the nature of ethics and the verifiability of ethical knowledge:

“No statement of fact can ever be, or imply, a judgment of absolute value. Suppose one of you were an omniscient person and therefore knew all the movements of all the bodies in the world dead or alive and that you also



knew all the states of mind of all human beings that ever lived, and suppose you wrote all you knew in a big book, then this book would contain the whole description of the world; and what I want to say is, that this book would contain nothing that we would call an ethical judgment or anything that would logically imply such a judgment.”

The second, future-oriented experiments are trying to presume what could happen in the future if something happens now, or if something that it is not possible now to happen, could happen and how it could affect on other future events. We can mentioned as, an example here, *The Case of Speluncean Explorers*, done by Lon Fuller in 1949, which was examining the behavior of five speluncean explorers trapped in a cave, and decisions of five judges, who conducted proceedings against the surviving explorers after their release from the cave. The specific of this thought experiment is that it happens in year 3400, that there are some rules of conduct that didn't exist at the time Fuller wrote it and that he was imagining everything- the plot, the rules and the solutions.

#### **4. Concluding remarks**

Thought experiments are product of the imagination. Yet they are important in every science, because their use enhances the knowledge in specific way. We find them in biology, economics, history, mathematics, philosophy, and physics. We find them in sociology, ethics, law. And every time, no matter how many times they've been discussed so far, we discuss about them with new energy, ideas and widened beliefs.

Are they mandatory instrument in gaining new knowledge? No, but they are rather fun, innovative and seems to be – never ending. Sometimes we can even save the world thinking about imagined scenarios and trying to predict people's behavior. But how can we learn about reality (if we can at all), just by thinking in such a way? We think that we can, only if we learn how to better perceive the world around us and listen to what it tells us. Using thought experiments, we can, maybe, achieve more effective form of mental perception and make bonds to other people and to our environment more efficient than we are doing it now.

Thought experiments are allowing us to open the doors of perception, imagination and think about solutions that are in our subconscious. We hope that this paper would open new doors to its readers and inspires them to explore hidden knowledge in ourselves, by using thought experiments.

**Ćorić Dragana**

Doktor nauka, docent, Pravni fakultet, Novi Sad, Univerzitet u Novom Sadu, Republika Srbija

## ZNAČAJ MISAONIH EKSPERIMENATA

**REZIME:** Čini se da je život u svetu pravnih normi lak: imate pravila ponašanja koja vas vode u mnogim životnim situacijama i šta će vam se dogoditi ako se ne budete ponašali kako je određeno. Pravne norme predviđaju budućnost, na izvestan način, i daju nam smernice za život. Iako pravni sistem svojim pravilima nastoji da pokrije sva područja društvenog života, postoje situacije koje se ne mogu predvideti u vreme donošenja te regulative. Ove praznine bi se mogle smanjiti usvajanjem naknadnih pravila ponašanja. Da bismo predvideli događaj koji se može dogoditi i predvideli ljudsko ponašanje u tim situacijama i ljudski odgovor na kaznu koja preti u slučaju kršenja pravila, poželjno je sprovesti – misaoni eksperiment.

Osnova misaonog eksperimenta može biti potpuno izmišljen, pa čak i trenutno nemoguć događaj, ili varijacija nekih poznatih i mogućih događaja. Ključno pitanje prilikom formulisanja misaonog eksperimenta je “šta ako”. Odgovori na ovo pitanje mogu početi sa “onda je moguće”, “tada će biti”, “može biti” ili nečim sličnim. Sadržaj odgovora će se razlikovati samo na osnovu vrednosti, verovanja i stavova onoga ko odgovara na pitanje “šta ako”. U našem radu ćemo ukratko predstaviti koncept misaonog eksperimenta, njegovu unutrašnju strukturu, vrste i, dajući neke primere, podstaći čitaoce da budu informisaniji o ovoj temi.

**Ključne reči:** *misaoni eksperiment, imaginacija.*

### References

1. Brown, J. R. & Fehige, Y. (2019). Thought Experiments. In: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Downloaded 2020, March 15 from <https://plato.stanford.edu/archives/win2019/entries/thought-experiment/>
2. Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*, New York and Oxford: Oxford University Press

3. Cole, D. (2020). The Chinese Room Argument. In: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Downloaded 2020, March 15 from <https://plato.stanford.edu/archives/spr2020/entries/chinese-room/>
4. Einhorn, H. J. & Hogarth, R. M. (1982). Prediction, Diagnosis, and Causal Thinking in Forecasting, *Journal of Forecasting*, 1(1), pp. 23–36
5. Fuller, Lon L. (1949). The Case of the Speluncean Explorers, *Harvard Law Review*, 62 (4), Downloaded: 2020, April 15 from <http://w.astro.berkeley.edu/~kalas/ethics/documents/introduction/fuller49.pdf>
6. Gedankenexperiment, *Merriem-Webster's Dictionary*, Downloaded 2020, March 15 from <https://www.merriam-webster.com/dictionary/gedankenexperiment>
7. Goodman, N. (1947). The Problem of Counterfactual Conditionals, *The Journal of Philosophy*, 44 (5), pp. 113–128
8. Kirk, R. (2019). Zombies, In: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Downloaded 2020, March 15 from <https://plato.stanford.edu/archives/spr2019/entries/zombies/>
9. Kuhn, S. (2019). Prisoner's Dilemma, In: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Downloaded 2020, March 15 from <https://plato.stanford.edu/archives/win2019/entries/prisoner-dilemma/>
10. Mach, E. (1960). *The Science of Mechanics*, 6th American ed. Translated by Thomas J. McCormack. La Salle: Open Court
11. Meynell, L. (2014). Imagination and insight: a new account of the content of thought experiments, *Synthese*, 191 (17), pp. 4149–4168
12. Putnam, H. (1973). Meaning and Reference, *Journal of Philosophy* 70, pp. 699–711
13. Putnam, H. (1985). The meaning of “meaning” (1975/1985). In: *Philosophical Papers. Vol. 2: Mind, Language and Reality*: Cambridge University Press
14. Rescher, N. (1991). Thought Experiment in Pre-Socratic Philosophy. In: Horowitz, T., & Massey, G.J. (eds.), *Thought Experiments in Science and Philosophy*, (pp.31–41), Rowman& Littlefield, (Savage)
15. Routley, R. (1973-1993). Is There a Need For a New, an Environmental, Ethic?, *Proceedings of the XVth World Congress of Philosophy*, Varna, 1: 205–10,1973; reprinted in M. Zimmerman et al. (eds.), *Environmental Philosophy: from Animal Rights to Radical Ecology*, Englewood, NJ: Prentice Hall: 1993, p.12–21
16. Sanna, L. J. (1998). Defensive Pessimism and Optimism: The Bitter-Sweet Influence of Mood on Performance and Prefactual and Counterfactual Thinking, *Cognition and Emotion*, 12 (5), pp. 635–665

17. Sarewitz, D. & Pielke, R. (1999). Prediction in Science and Policy, *Technology in Society*, 21(2), pp. 121–133
18. Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences* 3 (3), pp. 417–457
19. Searle, J. R. (1999). The Chinese Room, In: Wilson R.A. and Keil F. (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press
20. Wittgenstein, L. (1997). Lecture on Ethics, reprinted in : Darwall, Gibbard, and Railton, (eds), *Moral Discourse and Practice: Some Philosophical Approaches*, New York: Oxford University Press
21. Witt-Hansen, J. ( 1976). H.C. Ørsted, Immanuel Kant and the Thought Experiment, *Danish Yearbook of Philosophy*, 13, pp. 23–48
22. Yeates, L. B. (2004). *Thought Experimentation: A Cognitive Approach, Graduate Diploma in Arts (By Research): Dissertation*, University of New South Wales, Downloaded 2020, April 15, from: <https://ia803100.us.archive.org/4/items/TECA2004/TECA-%282004%29.pdf>