

Summary

Piotr Biłgorajski

Thought experiments in philosophy: origin, structure and functions

The dissertation aims at analyzing the origin, structure and functions of thought experiments in philosophy. The main problem of the dissertation is what thought experiments are and what significance they have for philosophy. This is reflected in the structure of the work, which consists of four chapters. The first two chapters, "The origin of thought experiments" and "Functions of thought experiments", present a general perspective on the history and application of thought experiments in various fields of philosophy. The last two chapters, "The imaginativeness of thought experiments" and "Counterfactual nature of thought experiments", are devoted to a detailed analysis of the essential features of thought experiments. Their aim is to answer the question about the importance of imagination in thought experiments and how thought experiments can be a tool for gaining knowledge about what is metaphysically possible.

The first chapter concerns the origin of thought experiments in philosophy. In this chapter I will examine the history of the term "thought experiment" and show how the understanding of this concept has changed in philosophy and in exact sciences (1.2). For further considerations, section 1.3, where I analyze the essential properties of the philosophical thought experiment, is pivotal. I claim that there is an analogy between thought experiments and real experiments, which manifests itself in the similarity of their structures: both procedures are created with a specific cognitive goal and consist in intentional changing of some state of affairs. However, the fundamental difference results from the nature of the state of things being changed: in a real experiment the material being changed is empirical and factual, while in a thought experiment it is imaginative and counterfactual. Therefore, in order to call a certain procedure a thought experiment, it should fulfill three conditions: (1) its base should be a certain counterfactual situation, (2) it should have a well-developed narration (reference to imagination), and (3) it should have a clearly defined cognitive goal. The adoption of such a characteristic allows to distinguish thought experiments from purely formal reasoning on the one hand, and on the other hand – from literature: thought experiments, contrary to most fictional works, are created with explicit cognitive intentions.

In section 1.4 I will discuss the most popular theories of thought experiments, according to which: (1) thought experiments are used to integrate available data into new conceptual

schemas, (2) thought experiments – through some kind of intuition – give us access to a Platonic world of necessary truths, (3) thought experiments are arguments, (4) thought experiments are a kind of real experiments, and (5) thought experiments as mental models of possible worlds.

The second chapter concerns functions of thought experiments. Using Karl Popper's typology I distinguish between critical and apologetic thought experiments. Critical (destructive) experiments are formulated against a theory and their aim is to point out contradictions in that theory or to undermine common intuitions, assumptions and conclusions of other thought experiments. The goal of constructive thought experiments is to provide examples that confirm a given theory. In part 2.2 I will show that the main function of thought experiments in philosophy is to provide counter-examples, which would undermine the necessary conditions for the applicability of concepts assumed by a given philosophical theory. In section 2.3 I will examine the relationship between intuition and thought experiments. I will show that the specificity of thought experiments is, on the one hand, to persuade the recipient to use his or her intuition to assess the soundness of the reasoning contained in the experiment, and, on the other hand, that thought experiments can serve as "intuition pumps", which means that their purpose is to induce certain intuitions in the recipient of the thought experiment. In part 2.4 I am going to analyze several selected classical philosophical thought experiments using the findings of the previous sections.

The third chapter concerns the imaginativeness of the thought experiment. The aim of this chapter is to prove that if imagination is one of the cognitive tools, thought experiments are a conscious, i.e. disciplined (methodical) use of this tool. In sections 3.2 and 3.3 I will discuss the problem of defining imagination and mental images. In section 3.4 – the problem of imaginative resistance, which, as I believe may be relevant to the evaluation of some thought experiments by providing a new way of criticizing them. Section 3.5 deals with the problem of cognitive functions of imagination. In section 3.6, I will set out the criteria for credible imaginations. In section 3.7, drawing on the views of Nancy Nersessian, I will present the characteristics of the narrative of a thought experiment. In section 3.8 I will discuss some similarities between thought experiments and literature, especially the *science fiction* genre. I argue that although thought experiments present in philosophy resemble in some respects ideas present in speculative literature, due to the criteria distinguished for thought experiments (counterfactual, narrative, clearly expressed cognitive intention), they cannot be identified with stories present in literature. In section 3.9 I will discuss the main objections to thought experiments, which are: (1) thought experiments are "intuition pumps" that create false intuitions rather than true beliefs, (2) philosophy should use science instead of *science fiction* or *fantasy*, (3) thought experiments appear where there is a lack of ability to formulate arguments clearly. I will argue that, contrary

to what critics claim, the colourful form of thought experiments on the one hand increases the attractiveness of philosophical content, and on the other hand thought experiments, contrary to "pure reasoning", may be open to new, original interpretations.

Chapter four deals with the counterfactual nature of thought experiments. I will propose to distinguish three levels of counterfactual situations described in the thought experiment: weak counterfactuality (situations that can be reproduced in reality), medium counterfactuality (situations that currently cannot be reproduced in reality, but perhaps they would be in the future, due to progress in science) and strong counterfactuality (situations that "in principle" cannot be reproduced in reality and that illustrate metaphysical theorems). I attempt to show that the key to understanding the importance of thought experiments in philosophy and, more specifically, experiments based on strongly counterfactual situations, is to explore the relationship between conceivability and imaginability, and possibility. Referring to the work of Frank Jackson, Stephen Yablo, George Bealer and David Chalmers, I will discuss the following issues: (1) what does it mean that something is conceivable, (2) what distinguishes conceivability from imaginability, (3) what type of possibility (logical, conceptual, nomological, metaphysical) conceivability would lead to, and (4) under what conditions would it lead to the conclusion that something is or is not possible. I then compare two competing concepts of conceivability: the concept of Yablo, who interprets conceivability in terms of imaginability, and of Chalmers, who presents it on the basis of so-called two-dimensional semantics.

In the last part of the fourth chapter I will analyze the "imaginative" concept of conceivability. Referring to the views of Peter van Inwagen, I will point out the weaknesses of Yablo's proposal. Although neither Yablo nor van Inwagen discuss thought experiments directly, I claim that their ideas can be developed to justify the usefulness of thought experiments. I try to show that if thought experiments are a variation of modal reasoning, the consequences of the findings on the nature of the relationship between conceivability and possibility will also apply to the importance of thought experiments in philosophy. I will argue that the function of strongly counterfactual thought experiments is, through the involvement of the imagination, to provide a reasonable presumption of the truthfulness of metaphysical theorems.